Real-time prediction of Litho-facies from drilling data using an Artificial Neural Network: A comparative field data study with optimizing algorithms

Article *in* Journal of Energy Resources Technology · June 2021

DOI: 10.1115/1.4051573

4 authors, including:

Amit Saxena
Rajiv Gandhi Institute of Petroleum Technology
31 PUBLICATIONS **68** CITATIONS

Journal of Energy Resources Technology. Received May 26, 2021;
Accepted manuscript posted June 23, 2021. doi:10.1115/1.4051573

Real-time prediction of Litho-facies from drilling data using an Artificial Neural Network: A comparative field data study with optimizing algorithms

Romy Agrawal[a], Aashish Malik[a], Robello Samuel[b], Amit Saxena[a]

[a]Rajiv Gandhi Institute of Petroleum Technology, Jais India

[b]Halliburton, Houston, TX, USA

Abstract

The lithology of the formation is known to affect the drilling operation. Litho-facies help in the quantification of the formation properties, which optimizes the drilling parameters. The proposed work uses the artificial neural network algorithm and an optimizer to develop a working model for predicting the lithology of any formation within the study area in real-time. The proposed model is trained using the formation data comprising 15-dependent variables from the Eagleford region of the United States of America. It builds a method for measuring or forecasting litho-facies in real-time when drilling through a formation. It uses general drilling parameters for better precision, including Rate of Penetration, Rotation per minute, Surface Torque, Differential Pressure, Gamma Ray Correlation, and a d-exponent correlation function. The proposed model compares and assesses various first-order optimization algorithm's efficiency, such as Adaptive Moment Estimation, Adaptive Gradient, Root Mean Square Propagation, and Stochastic Gradient Descent with traditional artificial neural network in quantitative litho-facies detection. The model can predict the complex lithology for vertical/inclined/horizontal wellbores in real-time, making it a novel algorithm in the industry. The developed algorithm illustrates an accuracy of 86 % using Adam optimizer when tested with the existing data and improves as the model is trained with more data.

Keyword: Litho-facies Prediction, artificial neural network, Drilling data, Drilling optimization

1. Introduction

The exploration initiatives in the oil industry result from tedious planning exercises. The field personnel collects the exploration data through various like 2D, 3D, and 4D seismic surveys. The analysis of the collected seismic data indicates the underground topology [1]. The gathered topology or formation data is inferred to get the litho-facies or the formation markers [2]. It preliminarily indicates the depth and thickness of each formation and performs the basic calculations for the mud weight, the volume, and the amount of cement needed. Drilling bit selection[3], lost circulation estimation [4], drilling parameter optimization [5], [6], shale problem

avoidance [7], differential pipe sticking prediction [8], and wellbore stabilization problems [9], [10] benefit from having prior knowledge on the formation to drilled. As the difficulty of the formations to be drilled increases, lithology estimation becomes increasingly essential for drilling bit optimization [11]. The loss of circulation is the most common and expensive drilling challenge that drilling engineers have often encountered. It costs oil producers much money, either because of the drilling fluid replacement or subsequent drilling issues like pipe sticking [12]. When active shale layers are present in the wellbore, wellbore instability is another significant concern that can be critical. Any of these issues has the potential to cost millions of dollars to resolve. While there are established close relationships between formation lithology and log data, well-defined expressions are impossible to obtain due to high nonlinearity and complication in their relationships[13]. Because of the uncertainties associated with reservoir measurements, recognizing subsurface litho-facies is a much-researched subject that remains a thought-provoking challenge.

Several methods have been proposed in the literature for predicting the formation type and lithology post and before the drilling operation [14]. The most common method is collecting the cuttings from shale shakers and analyzing them [15]. This method cannot determine the lithology at the time of drilling operation. It can only be effective when the cutting reaches the surface. Hence this method will always have a time lag in determining the formation properties proportional to the time taken by cuttings to reach the surface [16]. Traditionally, core analysis, geomechanical spectroscopy logs, Rock-Eval pyrolysis, and other qualitative methods have been used to identify subsurface litho-facies [17]. Several logging methods produce a wide variety of well logs for measuring the petrophysical properties of reservoir rock [18], including wireline, logging-while-drilling, measurement-while-logging [19]. Litho-facies and stratigraphic classification based on well log, core, and other data are being studied [20], [21]. However, the traditional approach is inconvenient, tiresome, and costly, requiring a high level of domain knowledge. Furthermore, only a few published studies using traditional methods allow for litho-facies identification while drilling and employ surface drilling parameters for this purpose.

With the recent advancement in technology and computing practices, people are familiar with artificial neural networks (ANN) and their unmatched capability to solve complex practical numerical problems [22]. ANNs comprise multiple nodes, which look like biological neurons of

Journal of Energy Resources Technology. Received May 26, 2021;
Accepted manuscript posted June 23, 2021. doi:10.1115/1.4051573

the human brain. The neurons interact with each other through a network of nodes. The nodes require input data to perform simple operations on it. Then, the neurons forward these results to other neurons. The output at each node is known as its activation or node value. The activation function decides the node value based on the input to activate it. In biological terms, the activation function is a function that checks the working of the neuron (decides whether the neuron is firing). Different activation functions are used in different layers depending on their accuracy [23]. The artificial neural network has been used more prominently in detecting litho-facies. In 2006, Qi and Carr identified carbonate litho-facies prevailing in the southwest Kansas field from wireline well log data by utilizing an artificial neural network (ANN) model [24]. In 2012, Wang and Carr used discriminant analysis, artificial neural networks (ANN), support vector machines (SVM), and fuzzy logic techniques to simulate the litho-facies of the Appalachian Basin in the United States [25]. They created a 3-D model of shale facies at the regional scale using core and seismic data and well logs. Unsupervised cluster analysis for litho-facies classification estimated the active sweet spots in the Barnett Shales as depicted by Avanzini and Balossino [26]. For mudstone facies found in the Mahantango-Marcellus and Bakken Shale, the USA, Bhattacharya and Carr compared the efficiency of unsupervised and supervised machine learning models [17]. Bhattacharya and Ghahfarokhi used SVM to identify shale litho-facies in the Bakken formation in North Dakota, USA [27].

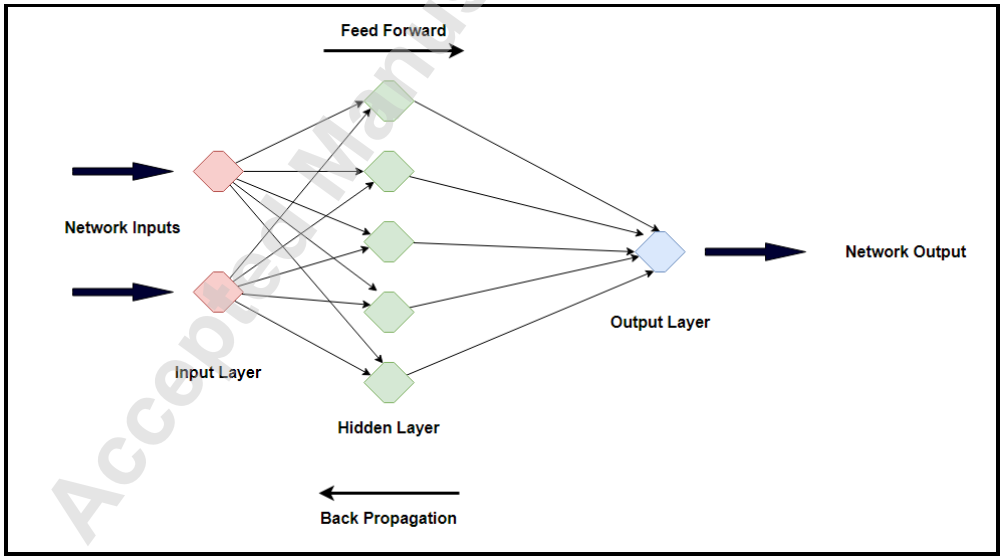


Figure 1. Working of Artificial Neural Network (ANN)

Journal of Energy Resources Technology. Received May 26, 2021;
Accepted manuscript posted June 23, 2021. doi:10.1115/1.4051573

85 **Table 1. Summary of some published research works related to litho-facies classification.**

86

S. No	References	Methods	Number of Input parameters
1	Qi and Carr (2006) [24]	ANNs	6
2	Al-Anazi and Gates (2010) [28]	Linear discriminant analysis (LDA), probabilistic neural network, general regression neural network and SVM	5
3	Raeesi et al. (2012) [29]	Competitive learning network and multilayer perceptron	2
4	Wang and Carr (2012) [25]	LDA, fuzzy logic, ANNs and SVM	5
5	Sebtosheikh and Salehi (2015) [30]	SVM	9
6	Horrocks et al. (2015) [31]	Naïve Bayes, ANNs and SVM	19
7	Avanzini et al. (2016) [26]	Ward's and K-means clustering algorithms	4
8	Bhattacharya et al. (2016) [17]	Self-organizing map, ANNs, SVM and multiresolution graph-based clustering	4
9	Al-Mudhafar (2017) [32]	Probabilistic neural networks	4
10	Tewari and Dwivedi (2018) [33]	Ensemble methods	15
11	He et al. (2019) [34]	Markov chain and ANNs	6
12	Gu et al. (2019) [35]	Probabilistic neural network, Boltzmann machine and particle swarm optimization	9
13	Imamverdiyev and Sukhostat (2019) [36]	Deep convolution neural network	7
14	Bhattacharya et al. (2019) [37]	SVM	5
15	Moradi et al. (2019) [38]	K-nearest neighbor, Naïve Bayes	7
16	Gupta et al. (2020) [39]	DTs, gradient boosting, random forest, and neural networks	10
17	Sun et al. (2020) [40]	Extreme Gradient Boosting (XGBoost), Bayesian Optimization (BO) and Gradient Tree Boosting-Differential Evolution (GTB-DE)	**7**

87

Journal of Energy Resources Technology. Received May 26, 2021;
Accepted manuscript posted June 23, 2021. doi:10.1115/1.4051573

Increased computing power has also resulted in more "big data," 3D litho-facies, and pattern stratigraphy applications. These annotated datasets combine vast collections of well logs and seismic data [37], [41]. Other algorithms for combining well-log analysis and automatic litho-facies classification that has recently gained traction include: Clustering using K-means [42], random forest [43], boosted generalized regression [44], and specific models of partitioning [45]. From the above literature, it is clear that conventional ANN approaches based on a multi-perception paradigm are helpful for the petroleum industry. However, recent data science research has led to new and more robust methods for optimizing ANN models. They use a variety of optimizers such as SGD (Stochastic Gradient Descent), ADAGRAD (Adaptive Gradient Algorithm), RMS PROP (Root-mean-square propagation), and ADAM. In ANN, different optimizing algorithms provide a strategic path to conceive the global minima by varying the weights. The optimizer minimizes the loss, and the system remembers the model parameters on which loss is minimum. The optimization function finds the global minima of the loss function. For a strictly convex function, the local minima equal the global minima. If the function is ostensibly convex, then the minima is estimated by speculating its neighborhood lowest value [46], [47].

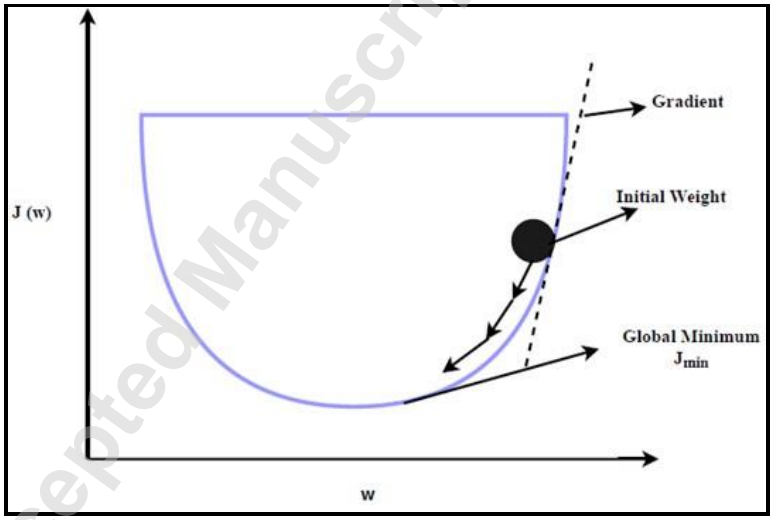


Figure 2. Minima of loss function from optimization function

The current work focuses on developing a procedure for determining or predicting the litho-facies in real-time while drilling through the formation. It uses general drilling parameters such as Rate of Penetration (ROP), (Rotation per minute) RPM, Surface Torque, Differential Pressure, Gamma Ray Correlation, and a correlation function d-exponent for better accuracy. The different

algorithms used in the study have enabled us to capture the patterns of the drilling parameters that affect the litho-facies and predict the litho-facies in the blind well. The other parameters used in the experimental procedure play a significant role as they help in optimizing the experimental result for better accuracy. This paper utilizes all four optimization methods listed above to evaluate the field data and conduct a comparative study for the validated data set. The significant contributions of this study are as follows:

- Analyzing the impact of extensive data analysis on the petroleum industry.
- Artificial neural networks (ANN) based models for quantitative litho-facies identification.
- Compare and assessment of the prediction efficiency of various optimizing algorithms in quantitative litho-facies detection.
- Comparison of traditional ANN models for testing and implementation of advanced ANN optimizers.
- Demonstrating a relationship between litho-facies and surface drilling parameters like ROP, RPM, Surface torque, etc.
- To provide for a means to detect the litho-facies in real-time while drilling.

This research will help reduce the drilling cost as it will provide an ultimate tool for predicting the subsurface litho-facies while drilling. In addition, it will help mitigate drilling problems such as lost circulation zones, differential sticking, etc.

2. Methodology

This paper focuses on predicting the litho-facies simultaneously while drilling, i.e., when the formation is drilled. Four wells from the Eagleford region of the USA are under consideration for the data. The collected data is sorted for drilling parameters and considered for the experimental procedure, which is assured by monitoring the depth. This means that data points with increasing depths are under consideration. It is evident from the literature that the input parameters for this experiment have been behaving differently in different formations [48]. The ANN model is trained using data from the three wells and tested on the fourth well. The total number of data points of three wells was 24,645, of which we used 75% of data points as training data and the remaining 25% as validation data. After training and validation of the model, it predicts the fourth well data comprising 8446 data points.

Journal of Energy Resources Technology. Received May 26, 2021;
Accepted manuscript posted June 23, 2021. doi:10.1115/1.4051573

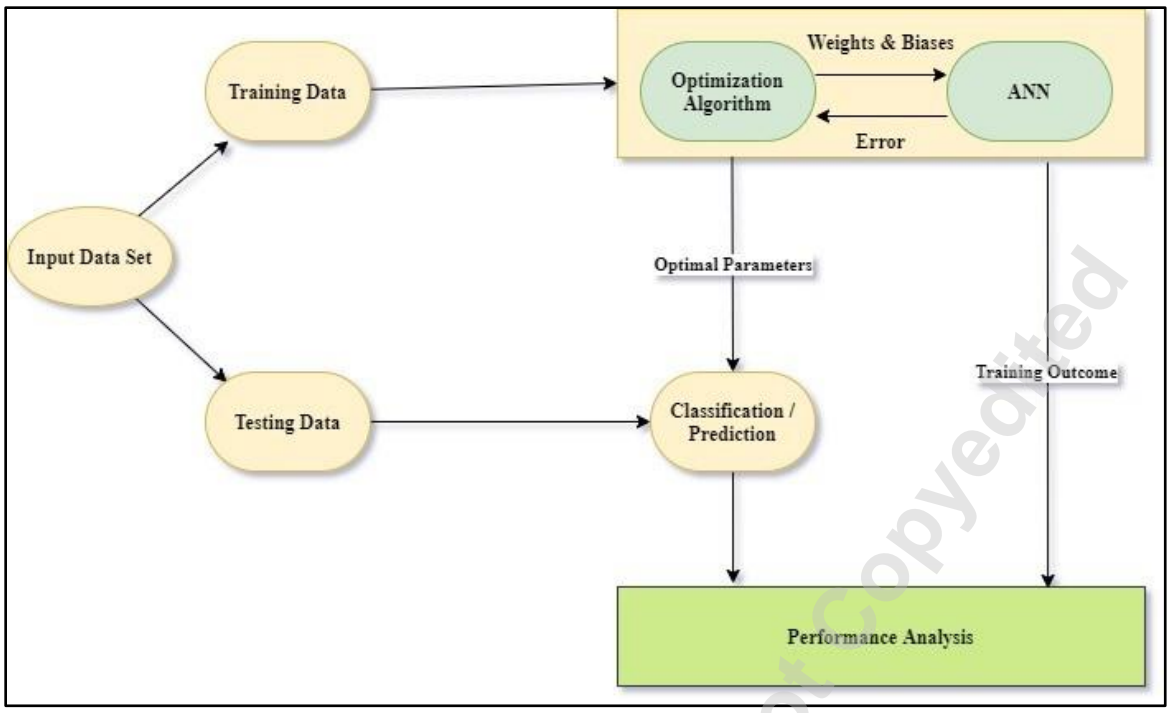


Figure 3. Flowchart of the Procedure used in the experiment

The model efficiency is predicted using two metrics, namely the F1 score and Accuracy Score. Where the F1 score is the weighted average of the precision and recall. F1 score reaches its best value at one and worst score at 0. The relative contribution of precision and recall to the F1 score are equal. Equation 1 represents the F1 score:

F1 = 2 * (precision * recall) / (precision + recall) Equation 1

Precision is the ratio of [tp/ (tp+fp)], tp is the number of true positives, and fp the number of false positives. The precision is intuitively the ability of the classifier not to label as positive a sample that is negative. The recall is the ratio $[t_p / (t_p + f_n)]$ where f_n the number of false negatives. The recall is intuitively the ability of the classifier to find all the positive samples.

In our case, the F1 score of each class was calculated and then combined by giving them weights based on the number of samples of each class [49].

Journal of Energy Resources Technology. Received May 26, 2021;
Accepted manuscript posted June 23, 2021. doi:10.1115/1.4051573

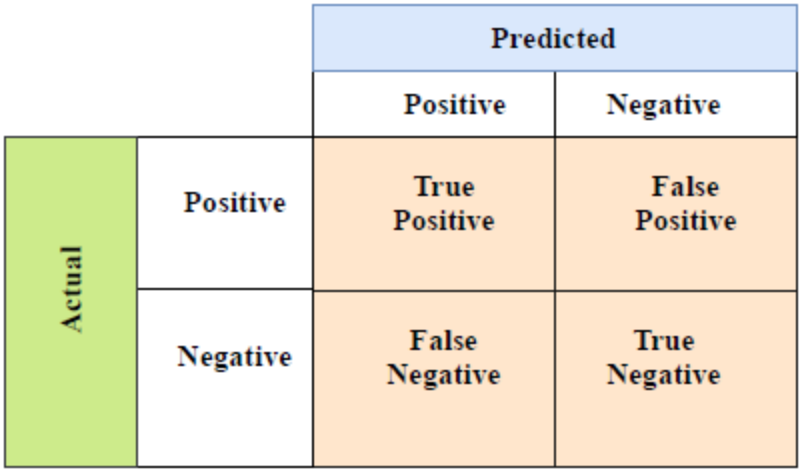


Figure 4. Confusion Matrix

The second metric used in our study is the Accuracy score, which is the ratio of total correct predictions to the total sample size. For example, If $\hat{y}_i$ is the predicted value of the i-th sample and y_i is the corresponding true value, then the fraction of correct predictions over $n_{samples}$ is defined as:

$$\text{Accuracy } (y,\hat{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}} 1\,(\hat{y}_i = y_i) \qquad \text{Equation 2}$$

where 1(y) is the indicator function.

The experiment uses 7 Neural network layers, comprising five hidden layers and the other being the input and the output layers. The number of nodes present in the hidden layers is:

a) First Layer – 87 Nodes
b) Second layer – 175 Nodes
c) Third Layer – 350 Nodes
d) Fourth Layer – 175 Nodes
e) Fifth Layer – 87 Nodes

Journal of Energy Resources Technology. Received May 26, 2021;
Accepted manuscript posted June 23, 2021. doi:10.1115/1.4051573

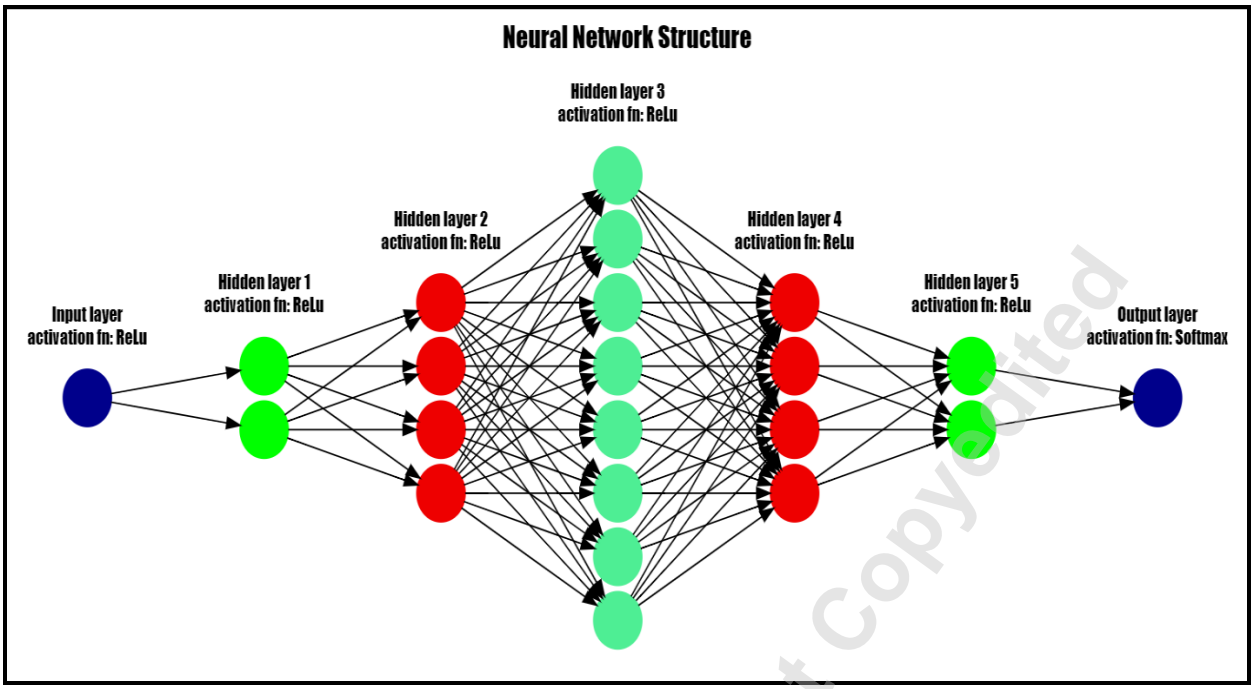


Figure 5. Neural Network structure used in the experiment

In this study, the ReLu function is used as an activation function in between the hidden layers, as shown in figure 5. It converts the values of the inputs which are less than zero, thereby forcing them to be zero and eliminating the vanishing gradient problem observed in other types of activation functions, and its formula is given by equation 3

$$y = \max(0, x) = f(x) = \begin{cases} x_i, & x_i \geq 0 \\ 0, & x_i < 0 \end{cases}, \quad \text{Equation 3}$$

The Softmax function converts a vector of 'n' real values into vector 'n' values whose sum is one. So, the softmax function transforms the input values into 0 to 1, which can be interpreted as probabilities [50].

The standard (unit) softmax function σ: $R^K \rightarrow R^K$ is defined by the formula

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \text{ for i= 1,....,K and } z = (z_1, ..., z_K) \in R^K , \quad \text{Equation 4}$$

K= No. of classes, z = outcome of each class from the ANN model before applying softmax function, and R stands for real values.

Softmax Loss or Categorical cross-entropy loss function is used to develop the proposed model [51], [52].

$$f(s)_i = \frac{e^{s_i}}{\sum_j^C e^{s_j}} \; ; \; CE = - \sum_i^C t_i \log(f(s)_i) , \quad \text{Equation 5}$$

Here $f(s)_i$ is a probability function and CE is the Cross entropy loss, t is the true value for that class of data point, and its value can be 0 or 1, C is the number of classes, s is the outcome of each class from the ANN model before applying the SoftMax function.

The number of layers and nodes in each layer was determined using cross-entropy loss on validation and training data in the present work. Initially, a single hidden layer with five nodes was utilized, with validation and training scores collected. Then the number of nodes was raised. The validation of architecture was done, and the training scores were recorded. The improved performance led to a subsequent increase in the number of nodes. The procedure was repeated, and one layer with 5 nodes was added, and nodes were raised until the validation and training performance. Finally, when adding more layers began to compromise model performance, the resultant architecture was finalized. The model architecture for our case has five hidden layers.

2.1 Input parameters:

In this experiment, 14 drilling parameters and one correlation parameter are under consideration for better accuracy of the ANN model. All of the input parameters used in this analysis have a specific association with the litho-facies, demonstrating this correlation is one of the primary goals of this article. ROP is a standard parameter for evaluating the drilling performance and the formation properties, such as lithology and hardness. ROP increases with increasing depth in formations such as sandstone, while it decreases in formations such as shale due to the rock's consolidation. ROP decreases in shale because of diagenesis and overburden stresses. Over pressured zones misleads and can give twice of ROP as expected, which indicates a well kick.

A formation has a unique Rate of Penetration (ROP) under specified borehole and operational conditions. Thus, ROP serves as one of the significant parameters for detecting litho-facies [53]. Drill string and heavy tubular weights force the bit against the bottom of the wellbore, resulting in cutting the rock underneath the bit. ROP increases linearly with an increase in WOB until the WOB value reaches the bit floundering point beyond which the ROP value decreases for any increase in WOB value. The increase in string weight is clubbed with an increase in RPM for an optimum increase in ROP value. RPM calculates the rotational speed with which the drill bit or the drill string rotates [54]. The rotary mechanism of the string causes the generation of a rotating force, known as torque, and it ultimately builds a hole.

Mud weight is a significant parameter in drilling as it controls the formation pressure and helps build a sufficiently stable wellbore. Drilling begins with a low mud weight, and it subsequently increases as the operation proceeds to a more considerable depth. The mud weight moderation provides significant benefits in mitigating the usual wellbore problems [55]. Exceptionally high mud weight leads to slower ROP because of chip hold down effect resulting in fracturing of formation leading to a complete loss of drilling fluid. The mudflow in rate is estimated from the suction tank, the last tank in the flow series. After this, the mud travels down the wellbore and comes upward through the annulus. The mud sample from the tank's outlet is with optimum concentration of all the additives added for handling downhole conditions and contamination. The mudflow out rate represents the amount of mud coming out of the well. It is estimated when the mud-carrying cutting reaches the surface before the solid removal [56]. Comparison between the properties of mudflow in and mudflow out sample helps study the properties of the formation. Another essential aspect derived from the effect of mud weight or mud gradient is the differential pressure. Differential pressure is the aftereffect of any excessive mud gradient with the depth, which affects the equivalent mud weight and hydrostatic pressure column [57]. This results in an increase of differential overbalance pressure making the hole cleaning operation a tedious operation. An increased bottom hole pressure can negatively impact ROP as the cuttings can be held at the bottom of the hole, increasing the frictional wearing of the drill bits [58]. It also decreases the hole cleaning efficiency of the drilling fluid. The hole cleaning efficiency of the drilling fluid can be enhanced by considering the strokes per minute or the pump discharge, which is the total drilling fluid discharge per minute available from the mud pumps on the drilling site. It is considered that the larger the pump discharge more significant is the efficiency of the system to lift the cuttings and hence would result in a faster ROP.

Bit size approximately represents the outer diameter of the hole being drilled. A large diameter bit generates a large volume of cutting, resulting in a slower ROP than a smaller size of bit. An appropriate size of the drill bit provides a stable and usable conduit for further drilling operation [59]. The proposed research takes into account a segment where the bit type and size do not change. The natural gamma radiation (GR) emitted by various naturally occurring radioactive elements in the formation is measured in real-time. The real-time gamma-ray correlation log helps quantify shaliness and well-to-well correlation and acts as a tool for detecting litho-facies more

effectively [11]. Other rig surface parameters such as hook load, block position, and pump pressure provide a better correlation in detecting predicted litho-facies.

D-exponent is a general correlation parameter that utilizes drilling parameters such as ROP, WOB, RPM, and bit size. The D-exponent plot can detect a shift from a normal pressure regime to an abnormal pressure regime. A shift in the D-exponent trend line indicates a well control event [60].

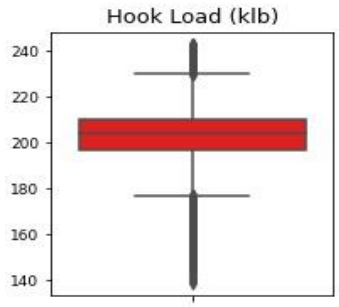
These parameters are used as a variable in training the proposed model. All of the wells investigated in this study have L–shaped profile. The model is trained to predict the complex geometries, including the well profile giving a momentary inclination over 90 degrees, i.e., bit returning into the previous strata for a smaller duration. We propose in this paper that wellbore inclination may be neglected as an input to the model, and the model can still predict formation litho-facies. The statistical parameter for the input data is illustrated in Table 2.

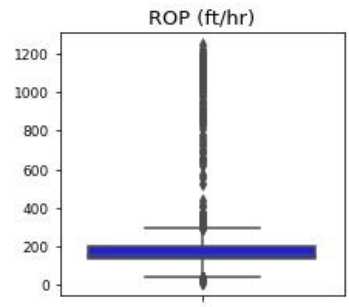
Table 2. Statistical parameters for the input data.

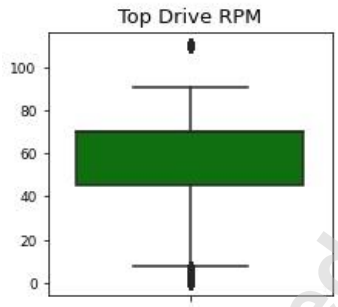
Statistical Parameter	HL (klb)	BT WT	BL HT	ROP (ft/hr)	RPM	TD Tor	D P (psi)	FL IN	SPM	FL OUT (%)	BT Size (inch)	GR (API)	Mud WT (lb/gal)	PP (psi)	D-EXP
Mean	203.4	26.2	43.83	183.8	60.29	5.9	516.2	589.6	179.4	36.9	8.75	86.3	10.19	2655.23	0.27
SD	11.4	10.	27.19	107.1	29.73	4.9	171.4	45.32	11.84	13.6	0	54.3	0.11	266.53	0.09
Min	139.1	0.10	-3.32	12.9	0.10	0	-1103	236	69	0	8.75	9	9.90	937.30	-0.38
25%	196.9	20	20.08	135.3	45	0	431.9	552	171	31	8.75	72	10.20	2520.50	0.25
Median	204.2	25.5	43.82	149.4	70	7.1	524.1	601	186	33	8.75	85	10.20	2658.77	0.30
75%	210.3	32	67.51	199.8	70.1	10.6	617.1	631	186	38	8.75	97	10.20	2821.19	0.33
Max	243.2	85.6	92.21	1248.2	110.6	15.6	1015.	632	186	93	8.7	858	10.4	3319.9	0.47
Kurtosis	2.04	2.3	-1.21	20.1	-0.49	-1.6	1.41	0.44	1.54	7.49	0	124	1.63	3.13	14.76
Skewness	-0.27	0.6	0	3.1	-0.68	-0.03	-0.57	-1.14	1.57	2.88	0	10.1	-0.50	0.94	-3.19

We have also plotted the selected input parameters using a box plot. This enables a better understanding of the distribution of data, as depicted in figure 6. In the plot, the middle line represents the median of data, and hence the different quartiles can be distinguished easily.

Journal of Energy Resources Technology. Received May 26, 2021;
Accepted manuscript posted June 23, 2021. doi:10.1115/1.4051573

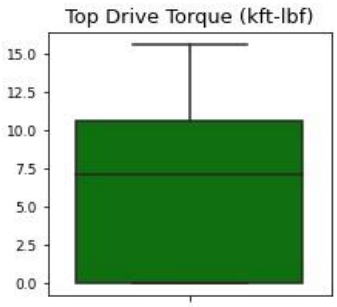


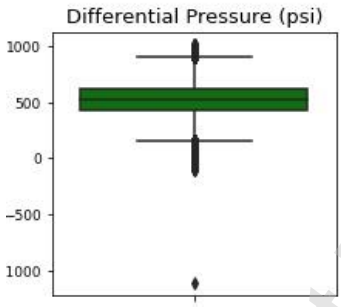


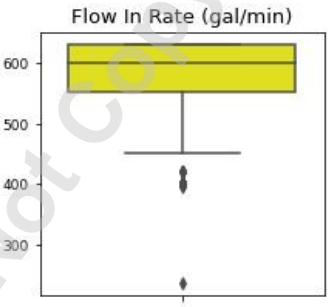


258

259 (6a) (6b) (6c)

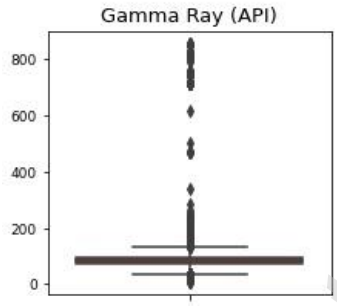


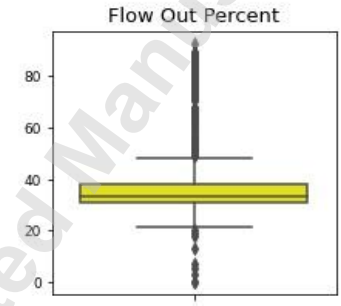


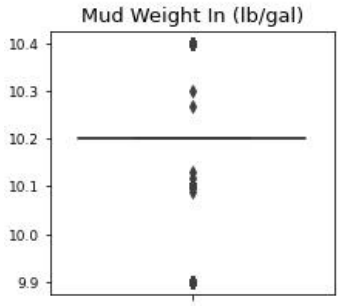


260

261 (6d) (6e) (6f)



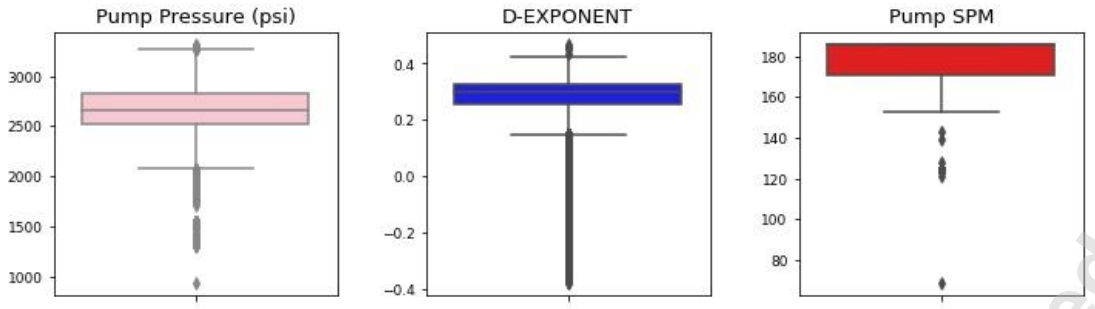




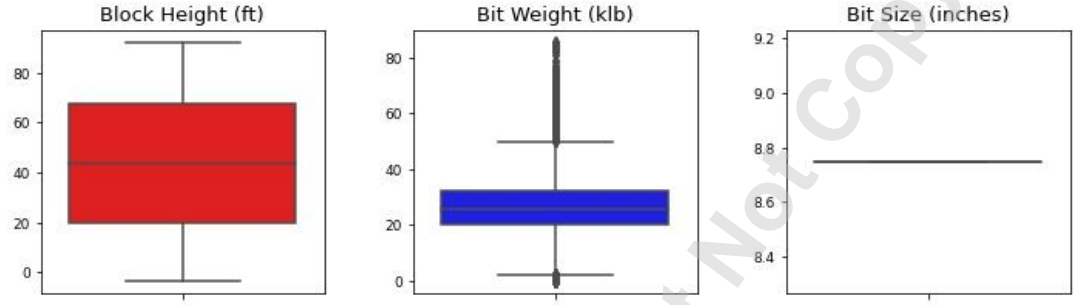
262

263 (6g) (6h) (6i)

Journal of Energy Resources Technology. Received May 26, 2021;
Accepted manuscript posted June 23, 2021. doi:10.1115/1.4051573



(6j) (6k) (6l)



(6m) (6n) (6o)

Figure 6. Box Plot of Input Parameters.

2.1.1 Outlier Removal:

This study utilizes an outlier removal technique (i.e., mean +3,-3 standard deviation technique) to select the outlier data point of each feature and verify each data point for at least two more features, discarding the same. The total number of outlier data points detected was 707 out of a total of 25352 data points utilized for training purposes (i.e., roughly 2.7 percent). Outlier removal will lead the model to converge at correct predictions as outlier values are spiked away from the general data distribution. The authors investigated all the features individually and discovered that the gamma-ray and hook load had a sensory error with a value of -999.25. The concentration of these error values was less than 0.5 percent prompting their elimination from the data set.

2.1.2 Dimensionality Reduction:

Journal of Energy Resources Technology. Received May 26, 2021;
Accepted manuscript posted June 23, 2021. doi:10.1115/1.4051573

The curse of dimensionality is a very well-known phenomenon that occurs while using artificial intelligence methods. A general approach to this is to remove unwanted features from the dataset. For this, the Pearson correlation test selected the highly correlated features. A maximum information coefficient (MIC) test on those highly correlated features can filter out the relevant features. Figure 7 shows the MIC score for each input parameter individually. The Pearson correlation test revealed that Pump SPM and Flow in Rate were strongly correlated in our situation. Then we ran a MIC test and discovered that flow in rate was more valuable, discarding the Pump SPM. Similarly, Block Height illustrated a meager MIC score leading to its elimination. The section had the same bit size, so its MIC score was coming out to be zero and was eliminated.

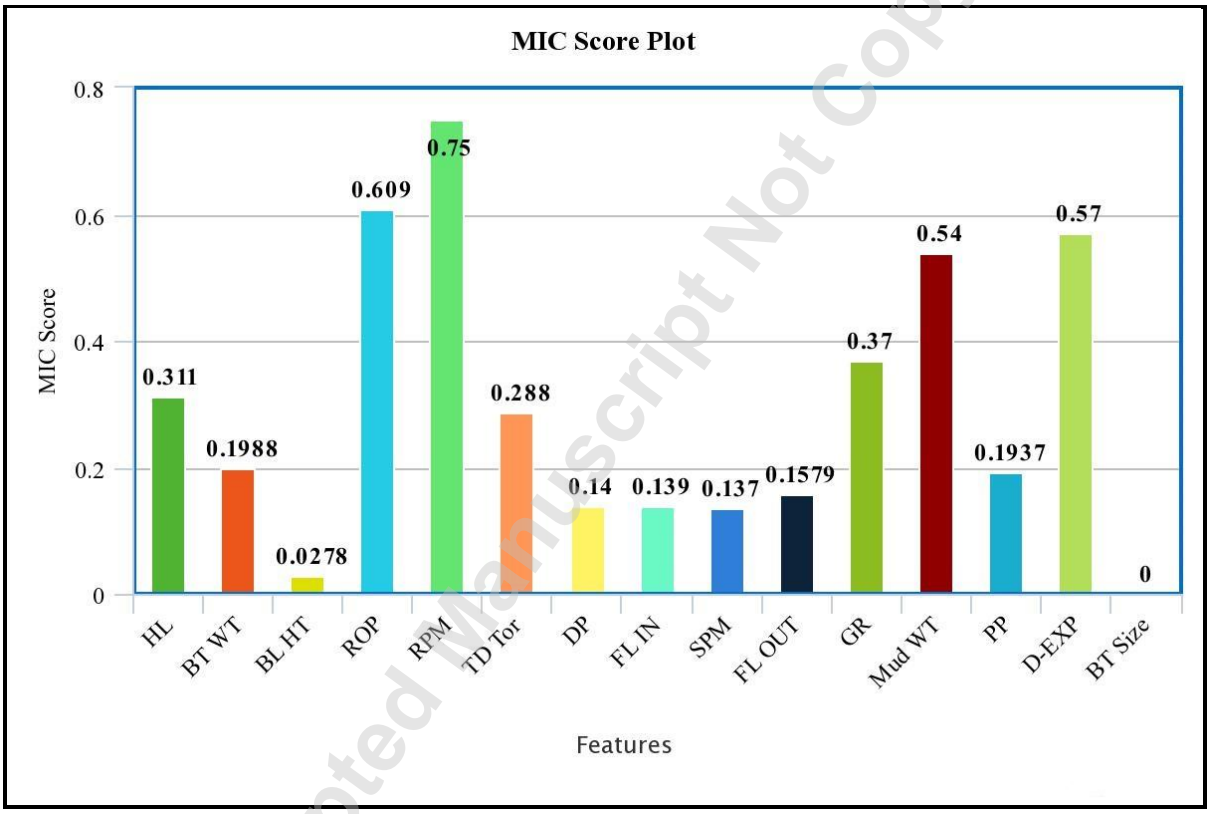


Figure 7. MIC Score Plot of the parameters under consideration.

2.2 Output Parameters:

The output of the ANN exercise performed over the given well data is the formation marker or the litho-facies as in such a manner that we have assigned a unique code to each formation type depicted in Table 3.

Journal of Energy Resources Technology. Received May 26, 2021;
Accepted manuscript posted June 23, 2021. doi:10.1115/1.4051573

294 **Table 3. The generated code for different lithology in the output**

Formation Name	Lithology	Code
Wilcox	Sand	0
Anacacho	Shale/Lime	1
Austin	Austin Chalk	2
Eagle Ford	Organic-Rich Marlstone	3
Buda	Limestone	4

295

296 From the available data, it is clear that two of our formations (Buda and Wilcox) account for
297 approximately 80% of the overall data, while Anacacho has the smallest number of data points.
298 As a result, we deal with data imbalance, and our paradigm would be skewed towards Buda and
299 Wilcox formations. The data imbalance problem is minimized using regularization parameters.

300 All hidden layers, as well as the input layer, were subjected to L2 regularization. The square of
301 the weights factor added to the error equation to punish the model for overfitting is known as L2
302 regularization. By keeping the weights and biases minimal, L2 regularisation attempts to limit
303 the likelihood of overfitting.

Journal of Energy Resources Technology. Received May 26, 2021;
Accepted manuscript posted June 23, 2021. doi:10.1115/1.4051573

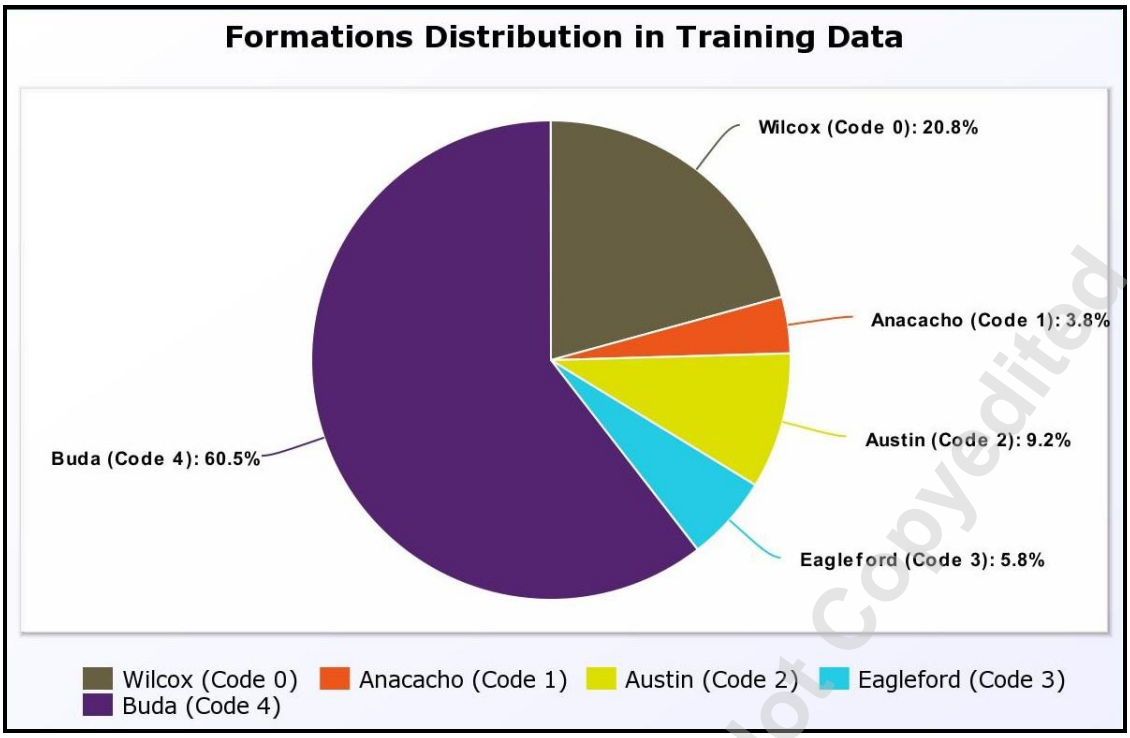


Figure 8. The plot of Formation distribution in Training data.

3. Result and Discussion:

Field data from four wells with over 31,000 data points were used to identify litho-facies in real-time while drilling. The input parameters and their significance in the drilling operation are described in section 2.1. The input parameters were trained accordingly to provide outputs for five different formations. The developed algorithm can be modified to predict many formations exhibiting a specific dependence over the drilling parameters. The proposed model was trained on three wells data (24645 data points, 75/25 ratio) with an Accuracy Score of 0.92 and an F1 score of 0.86. The developed model was tested on fourth (8446 unseen data points), and the results showed that Adam optimizer performed best with an Accuracy Score of 0.86 and F1 score of 0.81.

To validate the effectiveness of the ANN model, we have optimized the prediction results using four different optimizing algorithms (SGD, RMS PROP, ADAGRAD, ADAM). Among these algorithms, ADAM was found to be most efficient and promising. ADAM is one of the most refined and efficient algorithms used in ANN and deep learning. It can decide the individual weightage of the independent parameters and optimize the output parameters. In this research work, the traditional gradient descent-based ANN optimizer prominently used in the petroleum industry to identify the lithofacies is compared with other optimizers. The traditional optimizer

Journal of Energy Resources Technology. Received May 26, 2021;
Accepted manuscript posted June 23, 2021. doi:10.1115/1.4051573

performed better than ADAGRAD and RMS PROP but was poorer than ADAM and SGD, as illustrated in figure 9.

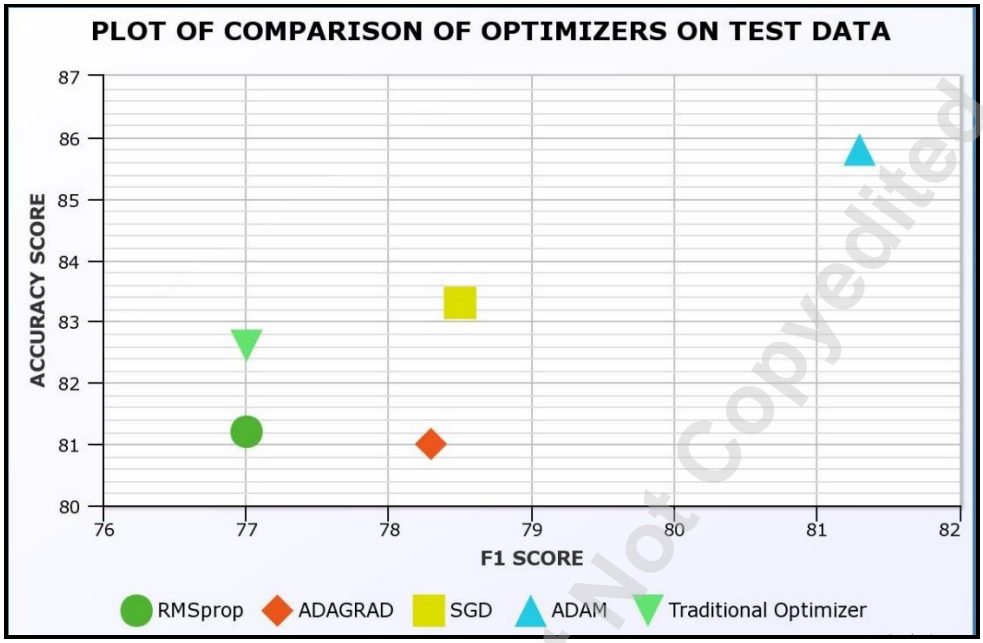


Figure 9. The plot of comparison of optimizers on test data.

The proposed work utilizes the confusion matrix to judge the model efficiency by observing the false positives, false negatives, true positives, and true negatives for each class (i.e., different lithology codes given in Table 2).

From figure 10 (a), i.e., the confusion matrix using Adam, we can observe that combining the overall percentage of each column will predict the overall data percentage of that formation. For example, in the first column, the overall predicted data percentage of formation code 0 is 20.51%, out of which 20.65% of data belongs to formation code 0. The rest falls in the category of false predictions. Similarly, the predicted and actual percentage of all the other remaining formations from the data and figure are represented in figure 10 (b to e). Within the formations, subsurface layers of varying thickness and patterns occur naturally. As a result, it has been assumed that an unequal data distribution reflects real-world conditions. This also allows one to understand the worst to the best performance of optimizers for individual layers under imbalanced data conditions. As can be seen, the issue of data imbalance is visible in the confusion matrix heat map. However,

Journal of Energy Resources Technology. Received May 26, 2021;
Accepted manuscript posted June 23, 2021. doi:10.1115/1.4051573

Adam has done a better job of predicting formations with low amounts of data, such as the Anacacho formation (code 1) and the Eagleford formation (code 3).

In contrast, other optimizers have failed to do so. It can also be seen that Adam performs well in loss generalization even with imbalanced data. At the same time, all other optimizers fail, as shown by the fact that the ADAGRAD model struggled in loss generalization and was unable to forecast formations with fewer data points, such as Eagleford. Data can be balanced to solve this data imbalance problem by creating synthetic data using techniques such as SMOTE, etc., but it is outside the reach of this paper.

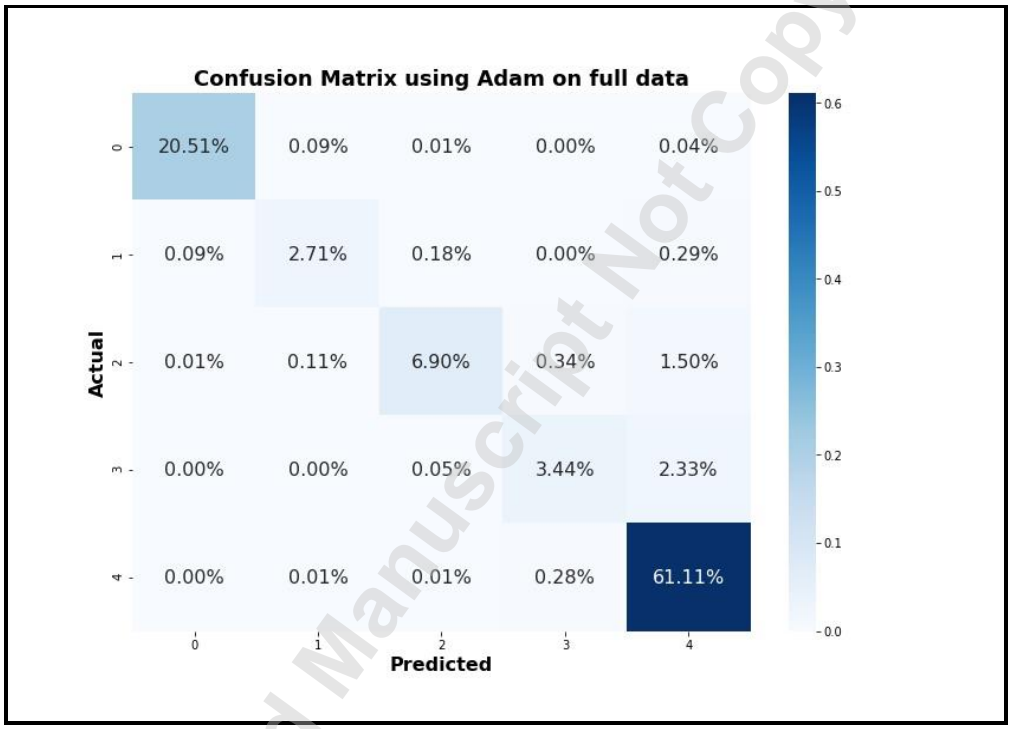
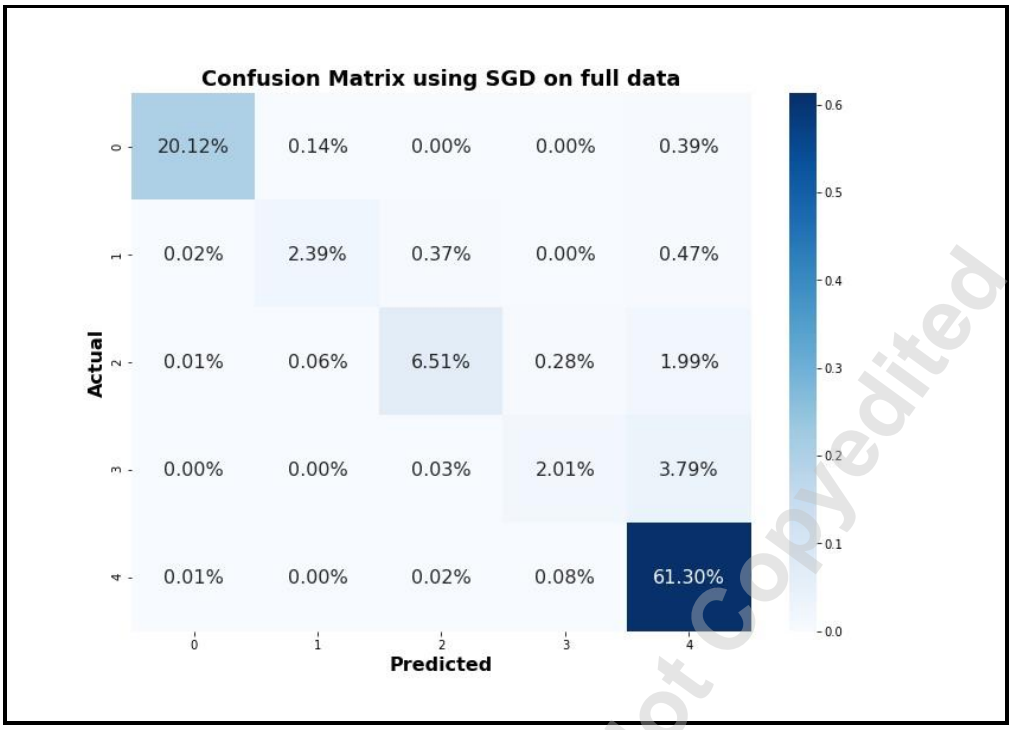


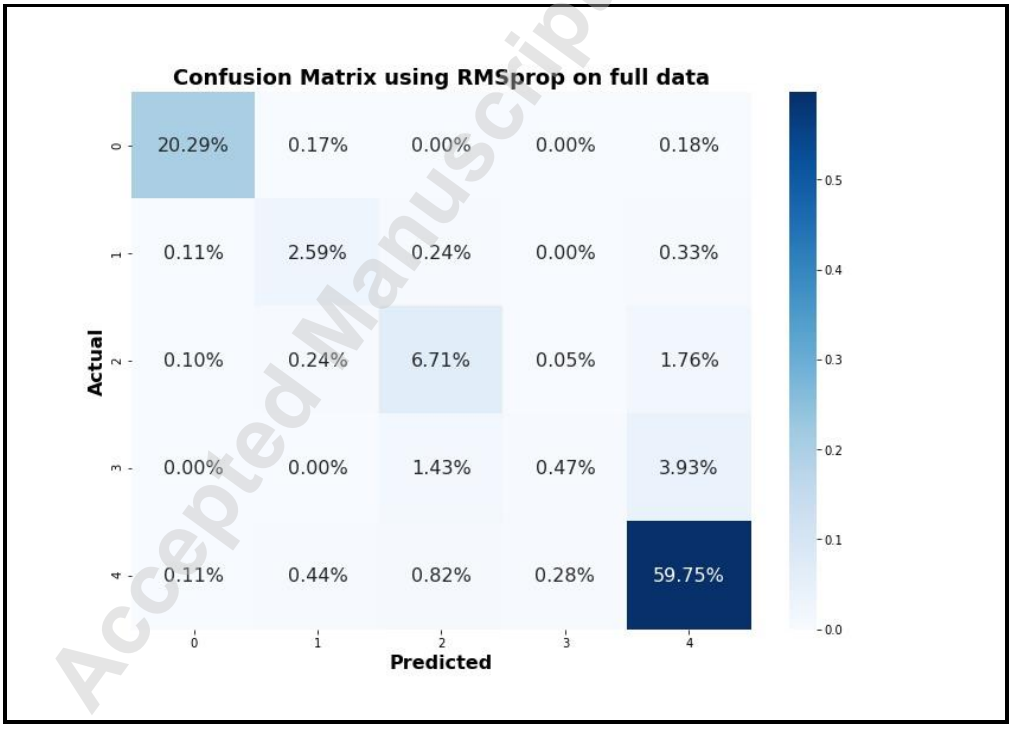
Figure 10 (a). Confusion matrix for Adam

Journal of Energy Resources Technology. Received May 26, 2021;
Accepted manuscript posted June 23, 2021. doi:10.1115/1.4051573



352

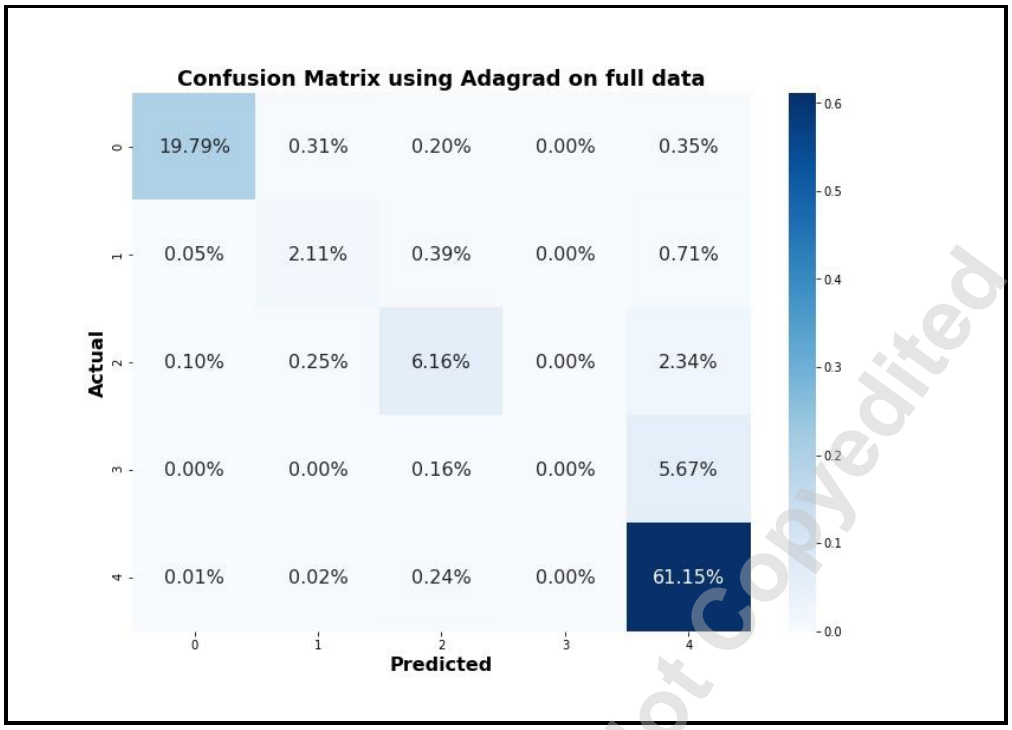
353 **Figure 10 (b). Confusion matrix for SGD.**



354

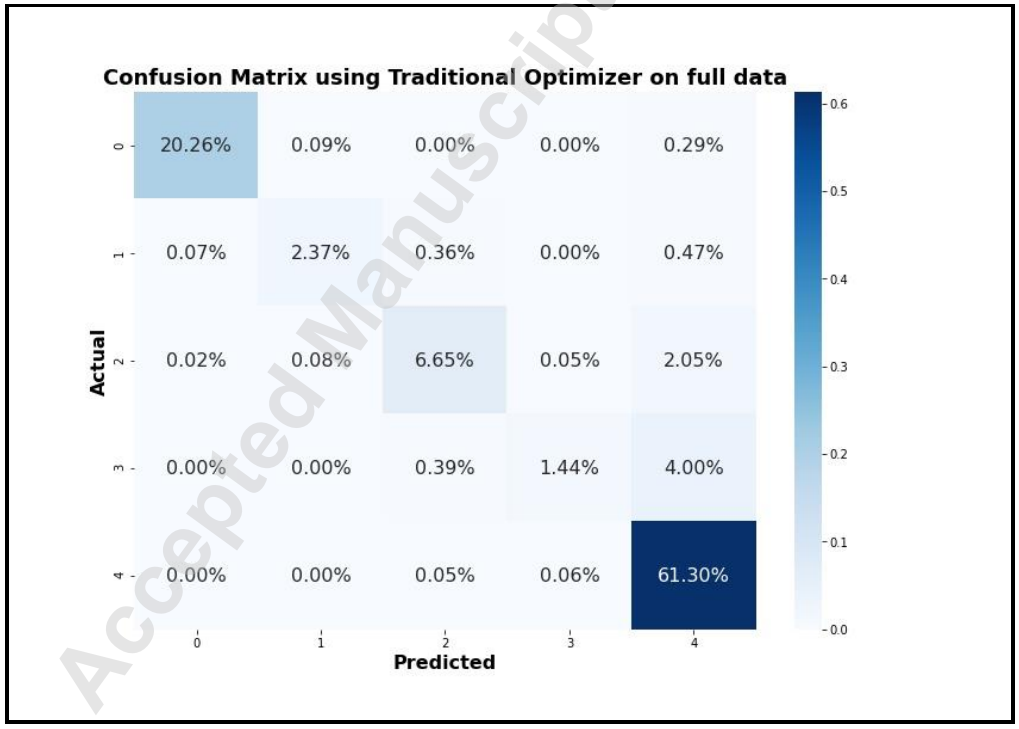
355 **Figure 10 (c). Confusion matrix for RMS PROP.**

Journal of Energy Resources Technology. Received May 26, 2021;
Accepted manuscript posted June 23, 2021. doi:10.1115/1.4051573



356

357 **Figure 10 (d). Confusion matrix for Adagrad.**



358

359 **Figure 10 (e). Confusion matrix for the traditional optimizer.**

Journal of Energy Resources Technology. Received May 26, 2021;
Accepted manuscript posted June 23, 2021. doi:10.1115/1.4051573

Torque and ROP are the two input features that impact the performance of models utilizing all four optimizers. When we compare it to domain knowledge, the impact of Torque and ROP on model performance is justified since these factors are crucial in deciding whether the formation is hard or soft. For example, while drilling a tight sandstone or shale section, the ROP is low, and the torque is high. While drilling a softer formation like carbonate, the ROP is high, and the torque is low. In contrast, the mudflow rate was shown to be the least important feature for all four models. The model's sensitivity was checked for the best performing optimizer, i.e., Adam, by removing the dominant parameters. It tabulated the top 5 of them to show the variation in accuracy score and F1 score results. For example, out of the 14 input parameters, we remove ROP, and the accuracy score of the model dropped to 0.827; at the same time, the F1 score will be 0.777. Similarly, the accuracy and F1 score have been reported for our study's five most dominant parameters. The results are elaborated in Table 3.

Table 3: Sensitivity Analysis of Input Parameters

Removed Parameter	Accuracy Score	F1Score
ROP	0.827	0.777
RPM	0.832	0.793
TORQUE	0.809	0.765
DP	0.836	0.790
GAMMA RAY	0.851	0.791
FINAL	0.86	0.81

The proposed research can be used to forecast litho-facies accurately. We plotted the actual formation against depth and compared it to the model's predicted formation. The lithology used here is complex as the formation with code three gets repeated after formation code 4. This is attributed to the fact that the actual data set is representing an inclined wellbore whose inclination exceeded 90 degrees momentarily. This causes the wellbore to re-enter the previous strata, making it difficult for the other models to predict the litho-facies. However, the same can be predicted by the proposed model. The proposed model clubbed with real-time data interpretation can provide a versatile tool for predicting the lithology of the unknown formation. The output formation number

will change as the number of subsurface strata varies. It will be more challenging for the model to distinguish a formation with fewer data points than one with many data points. If we find a reservoir with multiple thin formations, taking depth-wise data would result in fewer data points. As a rule of thumb, in reservoirs with multiple thin formations, it is recommended to use time-wise data, preferably 1 second time data, to maximize the number of data points.

The plot of actual and predicted formation against depth for the best optimizer (i.e., Adam in our case) has been depicted in figure 11. It can be observed that the majority of the variation lies in the section where formation changes. This fact holds accurate as while drilling, it is seen that when the bit drills in the transition zone, i.e. when the bit moves from one type of formation to another type of formation. This makes the prediction more complex in these overlapping sections.

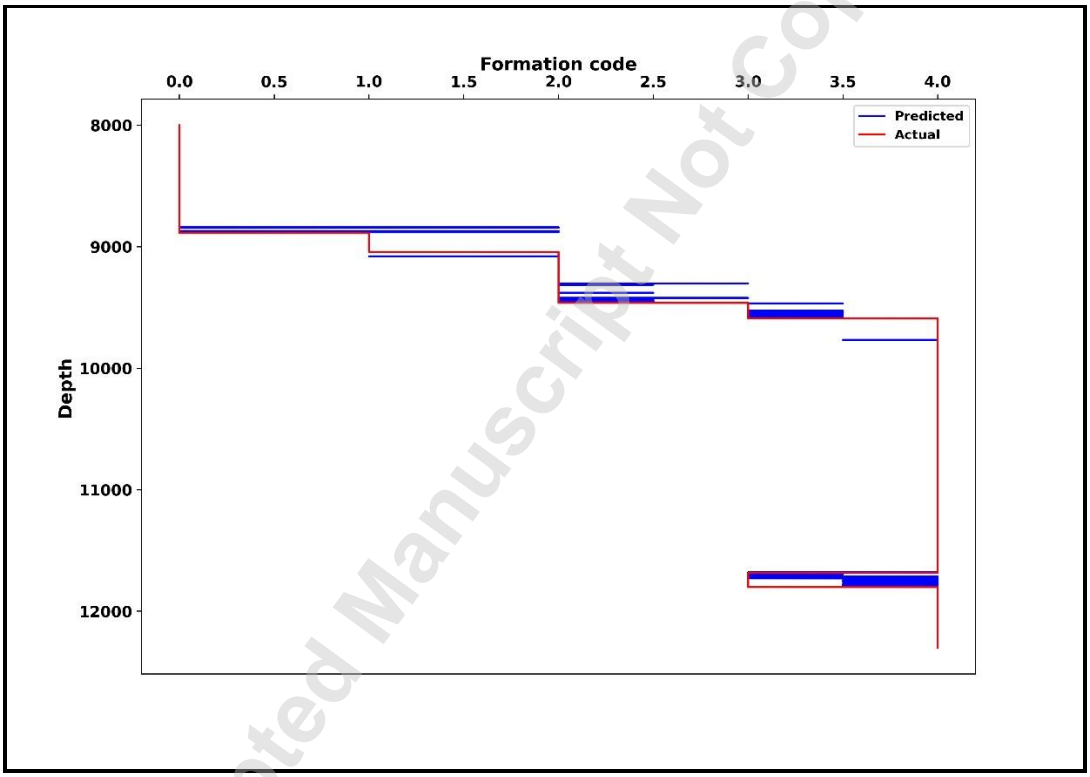


Figure 11. The plot of Predicated formation with depth for Adam optimizer

4. Conclusion

The development of novel ANN models is the need of the hour for the petroleum industry. It helps in the automation of different processes minimizing the uncertainty and the human effort. The proposed ANN model can predict the new formations accurately and be trained for different fields across different geological formations. A diligent formation-wise comparison has been made

between traditional ANN and newly developed optimizers to recognize litho-facies. In the prediction of litho-facies, Adam has an accuracy score of 86 percent and an F1 score of 81 percent, while the standard ANN model has an accuracy score of 82.6 percent and an F1 score of 77 percent. Compared to standard ANN the proposed model is 3.4 percent more accurate in terms of accuracy score and 4 percent more accurate in terms of F1 score. The proposed optimizer produced more stable models than the ones prevalent in the industry. Each of the ANN classifiers was developed using Eagleford area field data and proper hyperparameter optimization. The validation loss paired with training loss is frequently used for loss generalization to constrain the data imbalance effect. The findings can be bulleted as follows:

- ANN can be utilized very efficiently in the prediction of litho-facies.
- Drilling data and litho-facies have been discovered to have an important link. Torque and ROP, for example, are essential drilling parameters for forecasting litho-facies.
- New and advanced optimizers have shown promising results in extracting litho-facies information from real-time drilling data.
- The thin sections and the reoccurrence of litho-facies can be predicted easily by the proposed model.

As previously stated in this proposed work, the bit type and bit size employed in our study are the same throughout the section. The effect of varying drilling parameters while using other types and sizes has not been explored, which may be relevant in future research. Geometrical data from the wellbore, such as inclination and azimuth, might be used as input parameters to make a more robust model.

5. Abbreviation:

ANN: Artificial Neural Network

ROP: Rate of Penetration

RPM: Rotation per Minute

Adam: Adaptive Moment Estimation

ADAGRAD: Adaptive Gradient

RMSPROP: Root Mean Square Propagation

SGD: Stochastic Gradient Descent

2D: Two Dimensional

3D: Three Dimensional

4D: Four Dimensional

SVM: Support Vector Machine

USA: United States of America

ReLU: Rectified Linear Unit

CE: Cross Entropy Loss

WOB: Weight on Bit

GR: Gamma Radiation

MIC: Maximum Information Coefficient

SPM: Stokes per Minute

SMOTE: Synthetic Minority Oversampling Technique

HL: Hook Load

BT WT: Weight on Bit

BL HT: Block Height

TD Tor: Top Drive Torque

DP: Differential Pressure

FL IN: Mud Inflow Rate

FL OUT: Mud Outflow Rate

Mud WT: Mud Weight

PP: Pump Pressure

D-EXP: D-Exponent

BT Size: Bit Size

LDA: Linear discriminant analysis

DTs: Decision trees

XGBoost: Extreme gradient boosting

GTB-DE: Gradient Tree Boosting – Differential Evolution

BO: Bayesian Optimization

6. Acknowledgement

We greatly acknowledge Rajiv Gandhi Institute of Petroleum Technology for the computational and technical support. Thanks, are also extended to all the members associated with the work.

7. References

[1] K. Bjørlykke, "Petroleum geoscience: From sedimentary environments to rock physics," *Pet. Geosci. From Sediment. Environ. to Rock Phys.*, no. June 2014, pp. 1–508, 2010, doi: 10.1007/978-3-642-02332-3.

[2] L. Adeoti, O. Y. Adesanya, K. F. Oyedele, I. P. Afinotan, and A. Adekanle, "Lithology and fluid prediction from simultaneous seismic inversion over Sandfish field, Niger Delta, Nigeria," *Geosci. J.*, vol. 22, no. 1, pp. 155–169, 2018, doi: 10.1007/s12303-017-0018-4.

[3] I. G. Falconer, T. M. Burgess, and M. C. Sheppard, "Separating Bit and Lithology Effects from Drilling Mechanics Data," 1988.

[4] A. Moazzeni, M. Nabaei, and R. Kharrat, "A breakthrough in controlling lost circulation in a pay zone by optimizing the particle size distribution of shellfish and limestone chips," *Pet. Sci. Technol.*, vol. 30, no. 3, pp. 290–306, 2011, doi: 10.1080/10916466.2010.483438.

[5] P. Yi, A. Kumar, and R. Samuel, "Real-time rate of penetration optimization using the shuffled frog leaping algorithm," *J. Energy Resour. Technol. Trans. ASME*, vol. 137, no. 3, pp. 1–8, 2015, doi: 10.1115/1.4028696.

[6] R. Ashena, M. Rabiei, V. Rasouli, A. H. Mohammadi, and S. Mishani, "Drilling Parameters Optimization Using an Innovative Artificial Intelligence Model," *J. Energy Resour. Technol.*, vol. 143, no. 5, Feb. 2021, doi: 10.1115/1.4050050.

[7] A. N. Abugharara, B. Mohamed, C. Hurich, J. Molgaard, and S. D. Butt, "Experimental Investigation of the Effect of Shale Anisotropy Orientation on the Main Drilling Parameters Influencing Oriented Drilling Performance in Shale," *J. Energy Resour. Technol.*, vol. 141, no. 10, Apr. 2019, doi: 10.1115/1.4043435.

[8] R. Jahanbakhshi and R. Keshavarzi, "Intelligent Classifier Approach for Prediction and

Sensitivity Analysis of Differential Pipe Sticking: A Comparative Study," *J. Energy Resour. Technol. Trans. ASME*, vol. 138, no. 5, pp. 1–10, 2016, doi: 10.1115/1.4032831.

[9] H. Gamal, A. Alsaihati, S. Elkatatny, S. Haidary, and A. Abdulraheem, "Rock Strength Prediction in Real-Time While Drilling Employing Random Forest and Functional Network Techniques," *J. Energy Resour. Technol.*, vol. 143, no. 9, May 2021, doi: 10.1115/1.4050843.

[10] O. M. Siddig, S. F. Al-Afnan, S. M. Elkatatny, and A. Abdulraheem, "Drilling Data-Based Approach to Build a Continuous Static Elastic Moduli Profile Utilizing Artificial Intelligence Techniques," *J. Energy Resour. Technol.*, vol. 144, no. 2, May 2021, doi: 10.1115/1.4050960.

[11] K. Wang and L. Zhang, "Predicting formation lithology from log data by using a neural network," *Pet. Sci.*, vol. 5, no. 3, pp. 242–246, 2008, doi: 10.1007/s12182-008-0038-9.

[12] A. Moazzeni, M. Nabaei, and S. G. Jegarluei, "Decision making for reduction of nonproductive time through an integrated lost circulation prediction," *Pet. Sci. Technol.*, vol. 30, no. 20, pp. 2097–2107, 2012, doi: 10.1080/10916466.2010.495961.

[13] E. C. Onyia, "SPE 18166 Relationships Between Formation Strength, Drilling Strength, and Electric Log Properties," 1988.

[14] A. M. Mohammad Ali, "Artificial Intelligence for Lithology Identification through Real-Time Drilling Data," *J. Earth Sci. Clim. Change*, vol. 06, no. 03, pp. 3–6, 2015, doi: 10.4172/2157-7617.1000265.

[15] M. Karimi, "Drill-cuttings analysis for real-time problem diagnosis and drilling performance optimization," *Soc. Pet. Eng. - SPE Asia Pacific Oil Gas Conf. Exhib. APOGCE 2013 Maximising Matur. Elev. Young*, vol. 2, pp. 1295–1305, 2013, doi: 10.2118/165919-ms.

[16] S. Tiainen, H. King, C. Cubitt, E. Karalaus, T. Prater, and B. Willis, "Drill Cuttings Analysis—a New Approach To Reservoir Description and Characterisation; Examples From the Cooper Basin, Australia," *APPEA J.*, vol. 42, no. 1, p. 495, 2002, doi: 10.1071/aj01027.

[17] S. Bhattacharya, T. R. Carr, and M. Pal, "Comparison of supervised and unsupervised approaches for mudstone lithofacies classification: Case studies from the Bakken and Mahantango-Marcellus Shale, USA," *J. Nat. Gas Sci. Eng.*, vol. 33, pp. 1119–1133, 2016,

doi: 10.1016/j.jngse.2016.04.055.

[18] T. Moussa, S. Elkatatny, M. Mahmoud, and A. Abdulraheem, "Development of New Permeability Formulation from Well Log Data Using Artificial Intelligence Approaches," *J. Energy Resour. Technol. Trans. ASME*, vol. 140, no. 7, 2018, doi: 10.1115/1.4039270.

[19] F. Anifowose, J. Labadin, and A. Abdulraheem, "Improving the prediction of petroleum reservoir characterization with a stacked generalization ensemble model of support vector machines," *Appl. Soft Comput.*, vol. 26, pp. 483–496, 2015, doi: 10.1016/j.asoc.2014.10.017.

[20] C. Xu, Z. Heidari, and C. Torres-Verdín, "Rock classification in carbonate reservoirs based on static and dynamic petrophysical properties estimated from conventional well logs," *Proc. - SPE Annu. Tech. Conf. Exhib.*, vol. 5, pp. 3972–3986, 2012, doi: 10.2118/159991-ms.

[21] S. H. Lee and A. Datta-Gupta, "Electrofacies characterization and permeability predictions in carbonate reservoirs: Role of multivariate analysis and nonparametric regression," *Proc. - SPE Annu. Tech. Conf. Exhib.*, vol. OMEGA, pp. 409–421, 1999, doi: 10.2523/56658-ms.

[22] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. E. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," *Heliyon*, vol. 4, no. 11, p. e00938, 2018, doi: 10.1016/j.heliyon.2018.e00938.

[23] V. Sharma, S. Rai, and A. Dev, "A Comprehensive Study of Artificial Neural Networks," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 2, no. 10, pp. 278–284, 2012, [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.468.9353&rep=rep1&type=pdf.

[24] L. Qi and T. R. Carr, "Neural network prediction of carbonate lithofacies from well logs, Big Bow and Sand Arroyo Creek fields, Southwest Kansas," *Comput. Geosci.*, vol. 32, no. 7, pp. 947–964, 2006, doi: 10.1016/j.cageo.2005.10.020.

[25] G. Wang and T. R. Carr, "Methodology of organic-rich shale lithofacies identification and prediction: A case study from Marcellus Shale in the Appalachian basin," *Comput. Geosci.*, vol. 49, pp. 151–163, 2012, doi: 10.1016/j.cageo.2012.07.011.

[26] A. Avanzini, P. Balossino, M. Brignoli, E. Spelta, and C. Tarchiani, "Lithologic and geomechanical facies classification for sweet spot identification in gas shale reservoir,"

Interpretation, vol. 4, no. 3, pp. SL21–SL31, 2016, doi: 10.1190/int-2015-0199.1.

[27] S. Bhattacharya, P. K. Ghahfarokhi, T. R. Carr, and S. Pantaleone, "Application of predictive data analytics to model daily hydrocarbon production using petrophysical, geomechanical, fiber-optic, completions, and surface data: A case study from the Marcellus Shale, North America," *J. Pet. Sci. Eng.*, vol. 176, no. January, pp. 702–715, 2019, doi: 10.1016/j.petrol.2019.01.013.

[28] A. Al-Anazi and I. Gates, "A support vector machine algorithm to classify lithofacies and model permeability in heterogeneous reservoirs," *Eng. Geol.*, vol. 114, pp. 267–277, Aug. 2010, doi: 10.1016/j.enggeo.2010.05.005.

[29] M. Raeesi, A. Moradzadeh, F. Doulati Ardejani, and M. Rahimi, "Classification and identification of hydrocarbon reservoir lithofacies and their heterogeneity using seismic attributes, logs data and artificial neural networks," *J. Pet. Sci. Eng.*, vol. 82–83, pp. 151–165, 2012, doi: 10.1016/j.petrol.2012.01.012.

[30] M. Sebtosheikh and A. Salehi, "Lithology prediction by support vector classifiers using inverted seismic Attributes data and petrophysical Logs as a new approach and Investigation of training data set size effect on Its performance in a heterogeneous carbonate reservoir," *J. Pet. Sci. Eng.*, vol. 134, p. PETROL5323, Aug. 2015, doi: 10.1016/j.petrol.2015.08.001.

[31] T. Horrocks, E. J. Holden, and D. Wedge, "Evaluation of automated lithology classification architectures using highly-sampled wireline logs for coal exploration," *Comput. Geosci.*, vol. 83, pp. 209–218, 2015, doi: 10.1016/j.cageo.2015.07.013.

[32] W. J. Al-Mudhafar, "Integrating well log interpretations for lithofacies classification and permeability modeling through advanced machine learning algorithms," *J. Pet. Explor. Prod. Technol.*, vol. 7, no. 4, pp. 1023–1033, 2017, doi: 10.1007/s13202-017-0360-0.

[33] S. Tewari and U. D. Dwivedi, "Ensemble-based big data analytics of lithofacies for automatic development of petroleum reservoirs," *Comput. Ind. Eng.*, vol. 128, pp. 937–947, 2019, doi: https://doi.org/10.1016/j.cie.2018.08.018.

[34] J. He, A. D. La Croix, J. Wang, W. Ding, and J. R. Underschultz, "Using neural networks and the Markov Chain approach for facies analysis and prediction from well logs in the Precipice Sandstone and Evergreen Formation, Surat Basin, Australia," *Mar. Pet. Geol.*, vol. 101, no. July 2018, pp. 410–427, 2019, doi: 10.1016/j.marpetgeo.2018.12.022.

[35] Y. Gu, Z. Bao, X. Song, S. Patil, and K. Ling, "Complex lithology prediction using probabilistic neural network improved by continuous restricted Boltzmann machine and particle swarm optimization," *J. Pet. Sci. Eng.*, vol. 179, no. April, pp. 966–978, 2019, doi: 10.1016/j.petrol.2019.05.032.

[36] Y. Imamverdiyev and L. Sukhostat, "Lithological facies classification using deep convolutional neural network," *J. Pet. Sci. Eng.*, vol. 174, no. October 2018, pp. 216–228, 2019, doi: 10.1016/j.petrol.2018.11.023.

[37] S. Bhattacharya and T. R. Carr, "Integrated data-driven 3D shale lithofacies modeling of the Bakken Formation in the Williston basin, North Dakota, United States," *J. Pet. Sci. Eng.*, vol. 177, no. December 2018, pp. 1072–1086, 2019, doi: 10.1016/j.petrol.2019.02.036.

[38] M. Moradi, B. Tokhmechi, and P. Masoudi, "Inversion of well logs into rock types, lithofacies and environmental facies, using pattern recognition, a case study of carbonate Sarvak Formation," *Carbonates and Evaporites*, vol. 34, no. 2, pp. 335–347, 2019, doi: 10.1007/s13146-017-0388-8.

[39] I. Gupta *et al.*, "Looking ahead of the bit using surface drilling and petrophysical data: Machine-learning-based real-time geosteering in volve field," *SPE J.*, vol. 25, no. 2, pp. 990–1006, 2020, doi: 10.2118/199882-PA.

[40] Z. Sun, B. Jiang, X. Li, J. Li, and K. Xiao, "A data-driven approach for lithology identification based on parameter-optimized ensemble learning," *Energies*, vol. 13, no. 15, pp. 1–15, 2020, doi: 10.3390/en13153903.

[41] T. Zhao, V. Jayaram, A. Roy, and K. J. Marfurt, "A comparison of classification techniques for seismic facies recognition," *Interpretation*, vol. 3, no. 4, pp. SAE29–SAE58, 2015, doi: 10.1190/INT-2015-0044.1.

[42] W. J. M. Al-Mudhafar and M. A. Bondarenko, "Integrating K-means clustering analysis and generalized additive model for efficient reservoir characterization," *77th EAGE Conf. Exhib. 2015 Earth Sci. Energy Environ.*, no. June, pp. 2301–2306, 2015, doi: 10.3997/2214-4609.201413024.

[43] Y. Kim, R. Hardisty, E. Torres, and K. J. Marfurt, "Seismic facies classification using random forest algorithm," *2018 SEG Int. Expo. Annu. Meet. SEG 2018*, pp. 2161–2165, 2019, doi: 10.1190/segam2018-2998553.1.

[44] L. Zhang and C. Zhan, "Machine Learning in Rock Facies Classification: An Application of XGBoost," pp. 1371–1374, 2017, doi: 10.1190/igc2017-351.

[45] D. A. Wood, "Lithofacies and stratigraphy prediction methodology exploiting an optimized nearest-neighbour algorithm to mine well-log data," *Mar. Pet. Geol.*, vol. 110, no. May, pp. 347–367, 2019, doi: 10.1016/j.marpetgeo.2019.07.026.

[46] S. Sun, Z. Cao, H. Zhu, and J. Zhao, "A Survey of Optimization Methods from a Machine Learning Perspective," *IEEE Trans. Cybern.*, vol. 50, no. 8, pp. 3668–3681, 2020, doi: 10.1109/TCYB.2019.2950779.

[47] S. Ruder, "An overview of gradient descent optimization algorithms," pp. 1–14, 2016, [Online]. Available: http://arxiv.org/abs/1609.04747.

[48] S. Gstalder and J. Raynal, "Measurement of Some Mechanical Properties of Rocks And Their Relationship to Rock Drillability," *J. Pet. Technol.*, vol. 18, no. 08, pp. 991–996, 1966, doi: 10.2118/1463-pa.

[49] C. Goutte and E. Gaussier, "A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation," *Lect. Notes Comput. Sci.*, vol. 3408, no. April, pp. 345–359, 2005, doi: 10.1007/978-3-540-31865-1_25.

[50] C. E. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation functions: Comparison of trends in practice and research for deep learning," *arXiv*, pp. 1–20, 2018.

[51] Z. Zhang and M. R. Sabuncu, "Generalized cross entropy loss for training deep neural networks with noisy labels," *arXiv*, no. NeurIPS, 2018.

[52] K. Janocha and W. M. Czarnecki, "On loss functions for deep neural networks in classification," *Schedae Informaticae*, vol. 25, pp. 49–59, 2016, doi: 10.4467/20838476SI.16.004.6185.

[53] O. Ahmed, A. Adeniran, and A. Samsuri, "Rate of Penetration Prediction Utilizing Hydromechanical Specific Energy," *drilling*, 2018, doi: 10.5772/intechopen.76903.

[54] A. M. Abdul-Rani, K. Ibrahim, A. H. Ab Adzis, B. T. Maulianda, and M. N. Mat Asri, "Investigation on the effect of changing rotary speed and weight bit on PCD cutter wear," *J. Pet. Explor. Prod. Technol.*, vol. 10, no. 3, pp. 1063–1068, 2020, doi: 10.1007/s13202-019-00795-2.

[55] R. Rooki, "Estimation of Pressure Loss of Herschel–Bulkley Drilling Fluids During Horizontal Annulus Using Artificial Neural Network," *J. Dispers. Sci. Technol.*, vol. 36,

no. 2, pp. 161–169, 2015, doi: 10.1080/01932691.2014.904793.

[56] M. M. Razi, M. Mazidi, F. M. Razi, H. Aligolzadeh, and S. Niazi, "Artificial Neural Network Modeling of Plastic Viscosity, Yield Point, and Apparent Viscosity for Water-Based Drilling Fluids," *J. Dispers. Sci. Technol.*, vol. 34, no. 6, pp. 822–827, 2013, doi: 10.1080/01932691.2012.704746.

[57] J. Zare and S. R. Shadizadeh, "Managed pressure drilling to increase rate of penetration and reduce formation damage," *Pet. Sci. Technol.*, vol. 32, no. 15, pp. 1833–1842, 2014, doi: 10.1080/10916466.2010.540618.

[58] H. Y. Zhu, Q. Y. Liu, and T. Wang, "1324. Reducing the bottom-hole differential pressure by vortex and hydraulic jet methods," *J. Vibroengineering*, vol. 16, no. 5, pp. 2224–2249, 2014.

[59] A. Nascimento, D. Tamas Kutas, A. Elmgerbi, G. Thonhauser, and M. Hugo Mathias, "Mathematical Modeling Applied to Drilling Engineering: An Application of Bourgoyne and Young ROP Model to a Presalt Case Study," *Math. Probl. Eng.*, vol. 2015, no. 1960, 2015, doi: 10.1155/2015/631290.

[60] Y. P. Solano, R. Uribe, M. Frydman, N. F. Saavedra, and Z. H. Calderón, "A modified approach to predict pore pressure using the D exponent method: An example from the carbonera formation, Colombia," *CT y F - Ciencia, Tecnol. y Futur.*, vol. 3, no. 3, pp. 103–111, 2007.