Smart System for Predicting the Porosity and Permeability of Petro Physical Rock Types Using LSTM Algorithm

Promise Ezekiel
Department of Computer Science
Rivers State University
Port Harcourt, Nigeria
ezekielpromise27@gmail.com

Onate Taylor
Department of Computer Science
Rivers State University
Port Harcourt, Nigeria
taylor.onate@ust.edu.ng

Emmanuel Bennett
Department of Computer Science
Rivers State University
Port Harcourt, Nigeria
bennett.okoni@ust.edu.ng

***Abstract*—Porosity and permeability are critical factors in determining the productivity of an oil and gas well. Higher porosity and permeability values generally translate into higher production rates and longer well lifetimes. This paper presents a smart system for predicting the porosity and permeability of petro-physical-rock-types. The system model was developed and trained using only permeability and porosity data from a reservoir database. The model was trained using the four layers. The first layer contains an input neuron of 13 and is used relu as an activation function. The second layer contains a neuro of 64 and an activation function of tanh. The third layer contains an input neuron of 128 and an activation function of softmax, and finally, the fourth layer is the output layer that uses sigmoid as an activation function. Other hyperparameters used in training the model are loss= categorical_crossentropy, optimizer=adma, epoch, 20, and batch_size=32. The result obtained from the model shows an accuracy of 98.9% for training data and 99.1% for testing data.**

Keywords—Porosity, Permeability, Oil Reservoir, LSTM Architecture

I. Introduction

Porosity and permeability are crucial properties of rocks in oil and gas reservoirs. Porosity refers to the amount of empty space, or pores, in a rock, while permeability is the measure of how easily fluids can flow through these pores. In general, rocks with higher porosity and permeability are more desirable for oil and gas production because they allow for a better flow of hydrocarbons. The porosity and permeability of a rock depend on various factors, such as the type of rock, the size and shape of the pores, and the presence of natural fractures. For instance, sandstone rocks are typically more porous and permeable than shale rocks, which have a finer grain size and fewer interconnected pores [1].

Geologists use various techniques such as core analysis, well logging, and imaging methods to measure porosity. The core analysis involves extracting cylindrical rock samples from the reservoir and analyzing them in the laboratory. Well-logging uses sensors to measure the properties of the rock formation surrounding the wellbore. Imaging methods such as micro-CT scanning can provide high-resolution images of the pore space in a rock sample. Permeability, on the other hand, is more difficult to measure directly. In some cases, it can be estimated from porosity measurements, but in most cases, it requires a flow test using a fluid injection and production system. The flow test provides information on how easily fluids can flow through the rock formation and the rate at which they do so [2].

Porosity and permeability are critical factors in determining the productivity of an oil and gas well. Higher porosity and permeability values generally translate into higher production rates and longer well lifetimes. Low porosity and permeability can result in a lower production rate and limit the number of hydrocarbons in the reservoir.

One way to improve the porosity and permeability of a reservoir is through hydraulic fracturing. Hydraulic fracturing, or fracking, involves injecting a high-pressure fluid into the reservoir to create small fractures in the rock. These fractures increase the permeability of the rock and allow for better flow of hydra hydrocarbons to the wellbore. However, hydraulic fracturing can also have negative environmental impacts, such as water contamination and air pollution. As a result, there is increasing interest in developing alternative methods for improving the porosity and permeability of reservoirs [3].

One such method is called well logging. Well-logging is a technique that uses various tools to measure the properties of the rock formation surrounding a wellbore. These tools can measure porosity, permeability, and other important parameters such as lithology, fluid saturation, and pressure. Well-logging is a non-invasive method that can provide valuable information about the reservoir without needing core samples [4].

Well logging tools can be broadly classified into open-hole logs and cased-hole logs. Open-hole logs are used before the well is cased, and they measure the properties of the rock formation in its natural state. Cased-hole logs, on the other hand, are used after the well has been cased, and they measure the properties of the rock formation through the casing.

Various well logging tools are used to measure porosity and permeability, such as nuclear magnetic resonance (NMR), acoustic, and electrical logs. NMR logs use magnetic fields to measure the properties of the fluids in the pore space, which can be used to determine the porosity and permeability of the rock formation. Acoustic logs use sound waves to measure the speed of sound through the rock, which can also be used to determine porosity and permeability. Electrical logs measure the conductivity of the rock, which can be related to its porosity and permeability.

Well logging is an essential technique for the oil and gas industry, providing critical information for reservoir characterization, well placement, and production optimization. For instance, well logging can help identify areas of the reservoir with higher porosity and permeability, which are more likely to contain hydrocarbons. Well-logging can also help identify natural fractures and faults in the reservoir, which can affect the flow of hydrocarbons and the performance of the well.

Porosity and permeability are also crucial in enhanced oil recovery (EOR) techniques. EOR techniques recover additional oil from the reservoir after exhausting primary and secondary recovery methods. EOR techniques can include injecting gases, chemicals, or water into the reservoir to increase the mobility of the oil and improve its recovery—the wellbore. In water injection techniques, porosity and permeability are important to ensure that the injected water can displace the oil and push it toward the wellbore. Porosity and permeability can also affect the reservoir's ultimate recovery factor (URF). URF is the fraction of the original oil in place (OOIP) that can be recovered using various production methods. High porosity and permeability can result in a higher URF because they allow for a better flow of hydrocarbons to the wellbore. Conversely, low porosity and permeability can result in a lower URF and limit the number of recoverable hydrocarbons in the reservoir.

II. Literature Review

The models developed in [5] depended on routine primary survey data findings and details about the core depth, top, and base profundities of usable horizons. They compared various machine learning algorithms to settle on the most effective one. Together, the three stone characteristics were best predicted and generalised by a model with two hidden Neural network layers. Support Vector Machine and Linear Regression Algorithms, for example, nevertheless fared well on the dataset. As a whole, the method enables expecting the change of porosity and porousness during desalination in permeable rocks and measuring salt focus without direct estimations in a lab.

[6] used probabilistic neural networks (PNNs) to visualize lithofacies' successions as part of well-logging data to predict the transport of individual lithofacies at missing intervals. The good log considered neutron porosity, shale volume, and water immersion as depth indicators to characterize the lithofacies and show their permeability. Nevertheless, sand, sandstone, and permeability are the three distinct lithofacies types chosen for this study. The PNN successfully classified lithofacies precisely, as 95.81 percent of the expected lithofacies were correctly predicted.

The nano porous natural/mineral microstructure of the Marcellus Shale was the focus of a study by [1], who used two state-of-the-art rock 3D models to investigate the impact of repository imprisonment (weight) on the porosity and permeability of shales. The typical oil/gas supply throughout Marcellus Shale’s play's genesis time frame was replicated in five different pressure scenarios spanning depths of 2,000 to 6,000 feet. The advanced stone's pre- and post-pressure 3D models were analysed for their porosity and permeability. In both 3D models, the influence of weight on porosity and penetrability was relatively small. The trial results are relevant to evaluating the potential for shale supply creation under varying pressure settings and determining the oil-/gas setup.

[7] introduce a novel method for scaling multimodal porosity-penetrability relationships using multivariate underlying relapse in machine learning. To infer an upsized porosity-porousness relationship, we first isolated each sub-volumes primary attributes (porosity, stage availability, volume division, etc.) using image analysis tools. Then we relapsed these attributes against the tackled Darcy-Brinkman-Stokes (DBS) penetrability using an Extra-Trees relapse model. Following this, the AI-predicted up-scaled porosity-penetrability relationship was shown using a Darcy-scale stream and compared to full Darcy-Brinkman-STOKES (DBS) reproductions over ten tests with 3603 voxels. For all of the mathematical simulations, GeoChemFoam was used. The mathematical and AI upscaled models agreed well with the full Darcy-Brinkman-Stokes (DBS) recreations, with the AI model being significantly less computationally expensive.

To develop reliable penetrability and porosity prediction methods, [8] amassed 53 representative instances with logging data from the Tarim Basin's Lower Cambrian dolomite supply. The data from the centre and the logs were run through five normal porousness porosity relationships and six machine learning methodologies to evaluate their usefulness and predict how well the different methods would work. The R2 for their proposed coordinated approach to porousness prediction was 0.869, demonstrating that when the standard penetrability model was combined with Machine methods, the effect of heterogeneity on the prediction was mitigated.

In order to determine whether or not artificial neural networks and genetic algorithms can reliably forecast the eleventh of penetrability in near carbonate rocks, [9] presented a case ten research. For the twenty-one conditions with average porousness, the RMSE was lowest for the RGPZ condition, which had an RMSE of 220.458% when predicting the test dataset, and highest for the Berg condition, at 2.368%. The root-mean-square error (RMSE) for the genetic algorithm was 240.433%, and the RMSE for the artificial neural network algorithm was 0.38%, based on a correlation of 23. Its improved capability to exhibit readily available pore microstructures due to interdependent and competing diagenetic measurements was attributed to the superior execution of 25 machine learning algorithms over conventional approaches.

[10] offer a model for predicting oil supply porosity and penetrability using a machine learning concept and petrophysical data. Least Squares Support Vector Machine (LSSVM) was used, which is a comprehensive machine learning process. Northern Persian Gulf oil storage facility data, was used to design and test the Least Square Support Vector method. Researchers compared the Least Squares Support Vector Model's output to the relevant real-world petro data and the experimentally-obtained yields from various methods. It was found that the least squares SVM method normally has a total relative discrepancy of less than 1% between the methodology evaluations and the significant actual information.

[11] offer two novel connections for penetrability forecasting, one with and one without collecting analysis. The results are encouraging, as the mean score of 16 after 50 iterations are greater than 0.96. Using 19 fundamental patterns, the authors provided a substantial instrument for penetrability forecasts based on mathematical algorithms that recognise the pore structure .

400 fully saline solution-soaked 3D tiny CT images of Bentheimer and Clashach, miniature centre attachments, were used by [12]. Researchers used various three-measurement picture investigation techniques to determine the rock attributes (such as porosity and total penetrability) and extract pore structure data (such as pore throat conveyance, pore availability) pore harshness) from the images. The total image dataset is split in two for training and testing, with the first containing 80 percent of the images and the second 20 percent. A computer-based reasoning model checks and approves both

sets of results. The k-overlay approval approach is used to improve the precision of the model. Other data, such as rock mineralogy, might be fused with the produced model to increase its accuracy. Nonetheless, while initially limited to the aforementioned rocks, the developed model can be successfully expanded to other stone kinds if enough micro-CT images are available.

III. METHODS/METHODOLOGY

This paper aims to design a smart system for estimating the porosity and permeability of Carbonate Rocks Types. A detailed design of the proposed system can be seen in figure 1 below

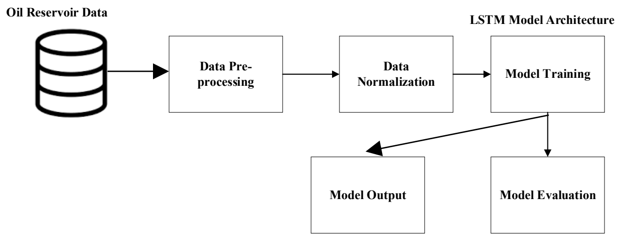

Figure 1: Architectural Design

A. Oil Reservoir Data

This study makes use of the very last set of data that was sampled from the Arab-D carbonate oil supply by Clerke. Currently, Clerke has accumulated nearly 450 High Pressure Mercury Injection Capillary Pressure (HPMI) estimations in the Arab D repository; however, his final examples were randomly chosen from 1,000's pre-qualified centre examples guaranteeing a comprehensive delivery and portrayal of all Petrophysical properties. Clerke used a Thomeer hyperbola fit to determine the bulk volume occupied (BVi), the shape of the Capillary Pressure bend associated with the inconsistency of the pore mouths (Gi), and the initial displacement pressures (Pdi) for each pore framework I in each sample. Clerke established the PRTs using this data, factoring in the number of pore structures and their corresponding Initial Displacement Pressures [13]. The first 10 rows of the dataset are displayed in Figure 2.

	Depth	PERMEABILITY	POROSITY	G1	PD1	G2	PD2	BV1	BV2	Mode	PRT	ROCK_INDEX	DATA_SOURCE
0	1	4800.00000	0.25810	0.49	1.29	0.15	1000.0	25.81	0.000	94.16	M	1	Rosetta
1	2	1550.00000	0.30050	0.69	1.55	0.15	1000.0	30.05	0.000	62.44	M	1	Rosetta
2	3	520.99350	0.30712	0.74	3.37	0.15	1000.0	30.71	0.000	27.12	M	1	Rosetta
3	4	234.00000	0.22080	0.64	4.39	0.15	1000.0	22.08	0.000	23.36	M	1	Rosetta
4	5	432.20240	0.24874	0.53	4.46	0.15	1000.0	24.87	0.000	26.20	M	1	Rosetta
...	...	...	...	...	...	...	...	...	...	...	...	...	...
439	440	0.00013	0.02057	0.25	8167.52	0.15	1000.0	2.06	0.001	0.02	3	6	Rosetta
440	441	0.06248	0.01387	0.22	10704.21	0.15	1000.0	1.39	0.001	0.02	3	6	Rosetta
441	442	0.00080	0.02062	0.34	11661.41	0.15	1000.0	2.06	0.001	0.01	3	6	Rosetta
442	443	0.00036	0.02185	0.05	12000.00	0.15	1000.0	2.18	0.001	0.02	3	6	Rosetta
443	444	0.00071	0.01889	0.23	15007.86	0.15	1000.0	1.89	0.001	0.01	3	6	Rosetta

Figure 2: Dataset Description

B. Data Pre-processing

The goal of data pre-processing is to prepare the data so that it is more easily understood by machine learning algorithms and thus improve the accuracy and reliability of the models. By properly pre-processing the data, we can ensure that the machine learning model can access high-quality data, leading to more accurate predictions and better performance.

C. Data Normalization

Normalization techniques can be applied to both porosity and permeability data individually or to the combined dataset. One commonly used normalization technique is Min-Max scaling, which scales the data between a minimum and maximum value. Another technique is z-score normalization, which standardizes the data to have a mean of zero and a standard deviation of one.

$$v_i' = (new\_max_A - new\_min_A) + new\_min_A \quad \text{Eqn. 1.}$$

Where A is the dataset features with n observed values $v_1, v_2, \ldots v_n$.

By normalizing the porosity and permeability data, we can ensure that the machine learning algorithm can use the data more effectively to make accurate predictions. Additionally, normalization can help prevent overfitting, which occurs when the model is too closely tuned to the training data and performs poorly on new data.

D. LSTM Model Architecture

An LSTM (Long Short-Term Memory) architecture was used to build a model for predicting porosity and permeability in the oil and gas sector. The following is the detailed architecture of the LSTM model:

1. **Input layer:** The input layer takes the sequence of data of input_shape=(13). The number of input nodes will depend on the length of the input sequence, and the size of each input node will depend on the dimensionality of the input data.
2. **LSTM layer:** This layer will have a certain number of LSTM units, allowing the model to learn the patterns in the input sequence over time. Each LSTM unit will have a cell state and a hidden state, which will be updated based on the current and previous input.
3. **Dense output layer:** This layer will have six output nodes, one for each of the six output classes. The activation function used for this layer is softmax, loss= categorical_crossentropy,

The mathematical equations of an LSTM architecture can be quite complex, but here are the main equations that describe how the LSTM units work:

At each time step t, given the input vector x_t and the previously hidden state h_{t-1}, an LSTM unit computes:

1. The forget gate f_t, which determines how much of the previous cell state to forget:

f_t = σ(W_f · [h_{t-1}, x_t] + b_f)

where σ is the sigmoid function, W_f and b_f are the weight matrix, and bias vector for the forget gate, and [h_{t-1}, x_t] represents the concatenation of the previous hidden state and the current input.

2. The input gate i_t, which determines how much of the current input to remember:

i_t = σ(W_i · [h_{t-1}, x_t] + b_i)

3. The candidate cell state ċ_t, which determines the new information to add to the cell state:

ċ_t = tanh(W_c · [h_{t-1}, x_t] + b_c)

4. The new cell state c_t, which combines the previous cell state with the new information:

$$c_t = f_t * c_{t-1} + i_t * \dot{c}_t$$

5. The output gate o_t, which determines how much of the cell state to output:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$$

6. The new hidden state h_t, which is a filtered version of the current cell state:

$$h_t = o_t * \tanh(c_t)$$

where W_i, W_f, W_c, and W_o are the weight matrices, and b_i, b_f, b_c, and b_o are the bias vectors for the input, forget, candidate, and output gates, respectively.

The LSTM architecture uses these equations to update the cell and hidden states at each time step, allowing it to learn patterns in the input sequence over time.

IV. Results

An experiment was carried out to build a smart system to predict porosity and permeability on carbonate rocks. The experiment phase comprises two phases. The first phase involves exploratory data analysis, and the second phase involves building an LSTM model to predict porosity and permeability in carbonate rock types. The following phases can be explained in detail:

A. Phase 1: Exploratory Data Analysis

We carried out an analysis of the dataset using charts. This was achieved using the matplotlib library in python. Figure 3 shows a correlation matrix of the various features on the dataset. The correlation matrix is used in checking if there exist relationships between the dataset attributes. A countplot that shows the different classes of the petrophysical rock types can be seen in Figure 4. A scatter plot that shows the different levels of porosity and permeability in carbonate rock types can be seen in Figure 5.

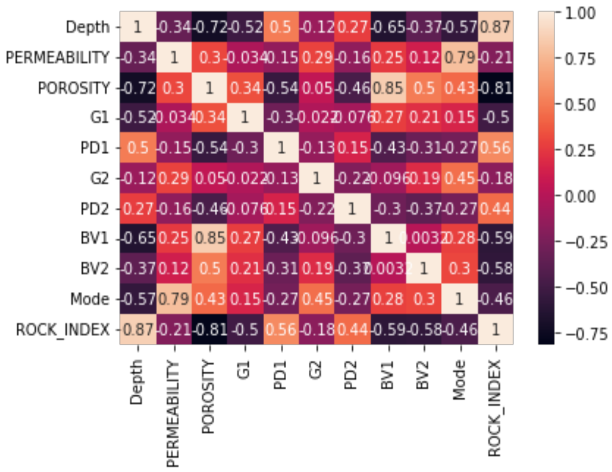

Figure 3: Correlation Matrix

The correlation matrix shows relationships between attributes in the dataset. The 1 in diagonal shows that there exist relationship features in the dataset.

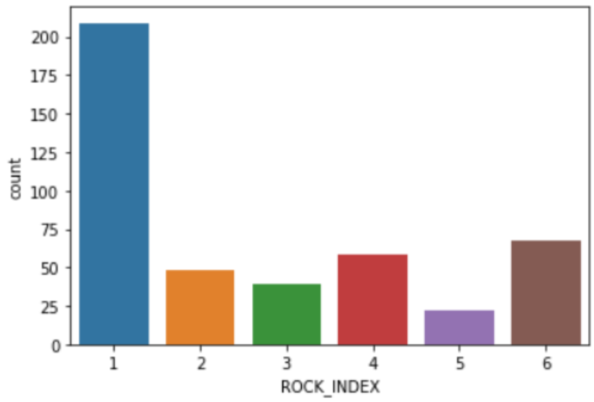

Figure 4: Countplot of the Rock index

The rock index represents the size of the pore system. The pore system determines how porous the carbonate rock is.

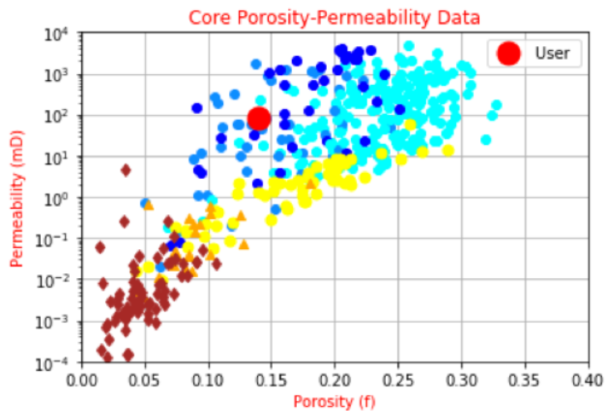

Figure 5: Core Porosity-Permeability data.

This shows a scatter plot of the porosity and permeability in the carbonate rock. The results show that the rock index changes at different permeability and porosity values. The rock can be either porous, meso porous, micro-porous, macroporous, etc.

B. Model Training with Lstm

The model was trained using LSTM. The model was trained using the four layers. The first layer contains an input neuron of 13 and is used relu as an activation function. The second layer contains a neuro of 64 and an activation function of tanh. The third layer contains an input neuron of 128 and an activation function of softmax, and finally, the fourth layer is the output layer that uses sigmoid as an activation function. Other hyperparameters used in training the model are loss= categorical_crossentropy, optimizer=adma, epoch, 20, and batch_size=32. The graphical representation of the loss and accuracy values during training and testing can be seen in Figures 6 and 7.

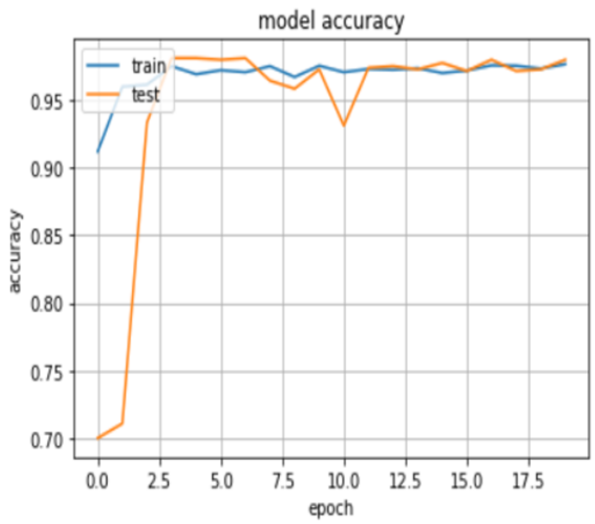

Figure 6: Model accuracy vs. Epoch

This shows the accuracy of the model at various training epochs. Here, the model achieved an accuracy of 98.9% for training data and 99.1% for testing data.

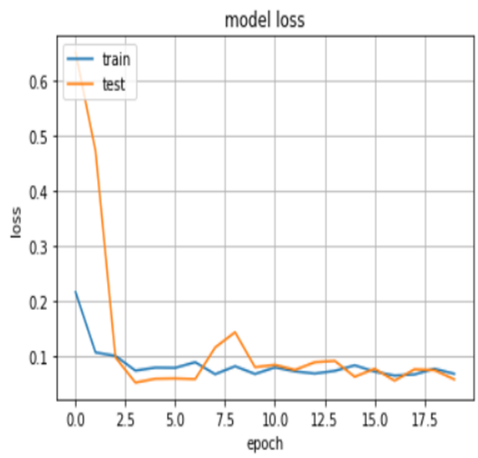

Figure 7: Loss vs Epoch

V Conclusion

This paper presents a smart system for predicting the porosity and permeability of petro-physical-rock-types. The system starts by training a model using just the permeability and porosity data obtained from a reservoir database. The model was trained using an LSTM model. The model was trained using LSTM. The model was trained using the four layers. The first layer contains an input neuron of 13 and is used relu as an activation function. The second layer contains a neuro of 64 and an activation function of tanh. The third layer contains an input neuron of 128 and an activation function of softmax, and finally, the fourth layer is the output layer that uses sigmoid as an activation function. Other hyperparameters used in training the model are loss= categorical_crossentropy, optimizer=adma, epoch, 20, and batch_size=32. This shows the accuracy of the model at various training epoch. Here, the model achieved an accuracy of 98.9% for training data and 99.1% for testing data.

References

[1]. P. S. Ezekiel, O. E. Taylor, and M. O. Musa, "A Framework for Predicting the Porosity and Permeability of Petro Physical Rock Types using Random Forest Classifier.", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 11, Issue 1, January 2022.

[2]. P. S. Ezekiel, O. E. Taylor and M. O. Musa, 2021. Smart System for Detecting Anomalies In Crude Oil Prices Using Long Short-Term Memory. International Journal, 10(6).

[3]. A.S. Al-Menhali, H.P Menke, M. Blunt, S.C. Krevor, “Pore scale observations of trapped CO2 in mixed-wet carbonate rock: Applications to storage in oil felds”. Environ. Sci. Technol. 50, 10282–10290 2016.

[4]. S.E. Kaczmarek, S.M. Fullmer, F.J.A. Hasiuk, “A universal classifcation scheme for the microcrystals that host limestone micro porosity”. J. Sediment. Res. 85, 1197-1212 2015.

[5]. A. Erofeev, D. Orlov, A. Ryzhov, D. Koroteev, “Prediction of Porosity and Permeability Alteration Based on Machine Learning Algorithms”, arXiv: 1902.06525v1 [cs.LG] 18 Feb 2019.

[6]. W.J. Al-Mudhafar “Integrating well log interpretations for lithofacies classification and permeability modeling through advanced machine learning algorithms”. J Petrol Explor Prod Technol 7, 1023–1033 (2017). https://doi.org/10.1007/s13202-017-0360-0.

[7]. H. P. Menke, J. Maes, S.Geiger, “Upscaling the porosity–permeability relationship of a microporous carbonate for Darcy scale fow with machine learning”, Scientific Reports 11(1), 1-10, 2021.

[8]. Z. Zhang, H. Zhang, J. Li, Z. Cai, “Permeability and porosity prediction using logging data in a heterogeneous dolomite reservoir: An integrated approach”, Journal of Natural Gas Science and Engineering, 86, 1-16, 2021.

[9]. H.A. Khalifah, p.w. Glover, P. Lorinczi, “Permeability Prediction and Diagenesis in Tight Carbonates Using Machine Learning Techniques”, Marine and Petroleum Geology, 112, 1-33, 2020.

[10]. M.A. Ahmadi, Z. Chen, “Comparison of machine learning methods for estimating permeability and porosity of oil reservoirs via petro-physical logs”. Petroleum 2018, doi:10.1016/j.petlm.2018.06.002.

[11]. Tran H., A. Kasha, A. Sakhaee-Pour, & i. Hussein, I, “Predicting carbonate formation permeability using machine learnin”g. Journal of Petroleum Science and Engineering,107581. doi:10.1016/j.petrol.2020.107581.

[12]. A.R. Shaik, A.A. Al-Ratrout, A.M. AlSumaiti, A.K. Jilani, “Rock Classification Based on Micro-CT Images using Machine Learning Techniques”. Abu Dhabi International Petroleum Exhibition & Conference. doi:10.2118/197651-ms.

[13]. E. Clerke, H.Mueller, W.Phillips, R.Wvazzadeh, D. Jones, R. Ramamoorthy, A. Srivastava, “Application of Thomeer Hyperbolas to decode the pore systems, facies and reservoir properties of the Upper Jurassic Arab D Limestone, Ghawar field, Saudi Arabia: A Rosetta Stone approach”, GeoArabia, Vol. 13, No. 4, p. 113-160, October, 2008.