

DOI: <http://doi.org/10.52716/jprs.v13i1.648>

Artificial Neural Networks to Predict Lost Circulation Zones at Southern Iraq Oilfield

Ameen K. Salih^{1,2,*}, Hassan A. Abdul Hussein², and Sarah H. Hamza³

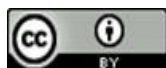
¹Petroleum Technology Department, University of Technology, Baghdad, Iraq

²Petroleum Department, Engineering College, Baghdad University, Baghdad, Iraq

³Petroleum Research & Development Center, Ministry of Oil, Baghdad, Iraq

*Corresponding Author E-mail: ameen.salih2008m@coeng.uobaghdad.edu.iq

Received 24/5/2022, Accepted in revised form 18/8/2022, Published 15/3/2023



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

Abstract

Drilling soft and fragile areas such as (high permeable, cavernous, fractured, and sandy formations) have several problems. One of the most critical problems is the loss of drilling fluid into these formations in whole or part of the well. The loss of drilling fluid can lead to more significant and complex problems, such as pipe sticking, well kick, and closing the well. The drilling muds are relatively expensive, especially oil-based mud or those that contain special additives, so it is not economically beneficial to waste and lose these muds. Artificial neural networks (ANN) can predict drilling fluid losses before they occur based on drilling parameters data and drilling fluid properties of wells effected by lost circulation problems located in the same area. This paper developed two artificial neural network models to predict drilling fluid losses in the Dammam formation- Rumaila oil field in southern Iraq. The two models have the same topology and structure. The first model used the early stopping technique to stop the training when we get the global minimum and the second model used specific epochs to complete the training. The models could predict various types of losses with high accuracy. The accuracy of implementing R2 for the first and second models was 0.9302 and 0.9493, respectively. The early stopping technique lead to obtain a model with acceptable accuracy in a short time without relying on a specific number of epochs.

Keywords: Artificial neural networks; ANN; Rumila oilfield; Early stop; Learning rate.

استخدام الشبكات العصبية الاصطناعية للتنبؤ بأحداث فقدان سائل الحفر في حقل نفطي جنوب العراق

الخلاصة

غالبًا ما تصاحب عملية حفر المناطق الرخوة والهشة مثل (التكوينات عالية النفاذية والتكوينات الكهفية ومناطق التصدع والرملية) العديد من المشكلات. ومن أهم هذه المشاكل فقدان مائع الحفر في هذه التكوينات كليًا أو جزئيًا. يمكن أن يؤدي فقدان مائع الحفر إلى مشاكل أكبر وأكثر تعقيدًا، بما في ذلك عسيان الأنابيب أو رفسة البئر وإغلاق البئر في النهاية. إن مائع الحفر غالي الثمن نسبيًا، خاصةً الطين الزيتي أو موائع الحفر التي تحتوي على إضافات خاصة، لذلك فهو غير مفيد اقتصاديًا إهدار هذه الموائع وفقدانها. توفر الشبكات العصبية الاصطناعية إمكانية التنبؤ بفقدان مائع الحفر قبل حدوثها بناءً على بيانات معاملات الحفر وخصائص مائع الحفر للآبار التي عانت من مشكلة الخسائر الموجودة في نفس المنطقة. في هذا البحث، تم تطوير نموذجين للشبكات العصبية الاصطناعية للتنبؤ بخسائر مائع الحفر في تشكيل الدمام - حقل الرميلة النفطي في جنوب العراق. النموذجان لهما نفس الهيكل والتركيب. استخدم النموذج الأول تقنية التوقف المبكر لإيقاف عملية التدريب عندما نحصل على الحد الأدنى للدالة دون الاعتماد على عدد محدد من التكرارات واستخدم النموذج الثاني عدد محدد من التكرارات لإكمال عملية التدريب. كانت النتائج متقاربة من حيث دقة النموذج وقدرته على التنبؤ بأنواع مختلفة من الخسائر. كانت دقة تنفيذ R^2 للنموذجين الأول والثاني 0.9302 و 0.9493 على التوالي. تقنية التوقف المبكر تمكنا من الحصول على موديل ذو دقة مقبولة بوقت قصير جدا دون الاعتماد على عدد محدد من التكرارات.

1. Introduction

The decreased volume of drilling fluid returning to the surface after cooling the bit, carrying the rock cuttings, and building the mud cake facing the permeable zones is called lost circulation [1]. The layers in which the loss of drilling fluid occurs are called thief zones [2]. Circulation loss occurs in highly permeable, depleted, natural fissures, cavernous, and fracture formations [3, 4]. The loss of drilling mud categorization depends on losses rate, which ranges from seepage losses ($0.16 - 1.6\text{m}^3/\text{hr}$), partial losses ($1.6-80\text{m}^3/\text{hr}$), and finally, complete or total losses when the losses more than ($80\text{m}^3/\text{hr}$) [5]. This problem is one of the costliest problems in the oil industry, as it costs \$2 billion annually to treat; it also represents (12% of NPT) according to worldwide oil industry estimation and (46% of NPT) in the Rumaila oil field [6].

The first step in dealing with drilling fluid losses is to determine the type of losses; if the losses are seepage or partial losses, LCM (lost circulation materials) pills are a good choice, and in case of severe or total losses, cement squeeze is a suitable choice [7]. The cost of treating drilling fluid losses is due to the cost of the materials used in addition to the cost of NPT such as (Cement, DOB, DOBC plug) need (18, 10, 12) hours and cost about (27, 15, 18) thousand USD respectively [8].

There are many factors affecting the loss of drilling fluid, including the petrophysical properties of the rocks (porosity, permeability, etc.) and the properties of the drilling fluid itself (MW, ECD, YP, PV, etc.) as well as the drilling parameters (ROP, WOB, RPM, SPM, SSP, TFA, etc.). In addition to (pore pressure gradient, fracture pressure gradient, etc.), these are some well-known factors and other unknown factors [9]. Controlling these factors to prevent the loss of drilling fluid is a challenging task, so it is necessary to have an intelligent model to predict the occurrence of losses or not. Predict the type of those losses depending on these factors. Therefore, some of these factors can be controlled to prevent or reduce the loss of drilling fluid [10]. Artificial neural networks are one of the most important techniques used in solving complex problems by revealing patterns and complex relationships between the causes of the problem and the outcome [11]. This study developed two artificial neural network models to predict circulation loss in the Rumaila oil field.

1.1 Artificial Neural Networks (ANNs)

Artificial Neural Networks are a subset of machine learning and the heart of deep learning algorithms. Neural networks are non-algorithmic, distributive, and parallel information processing analog methods that have proven to be powerful pattern recognition tools [12]. So, it can recognize highly complex relationships between the inputs and the outputs. A neural network is a computational model that includes many neurons and connections, and each can receive signals through synapses. When the signal is strong, it will activate the neuron, and the signal will go through the axon away to another neuron, and the new neuron may be activated or not depending on the strength of the signal [13], as shown in Figure (1).

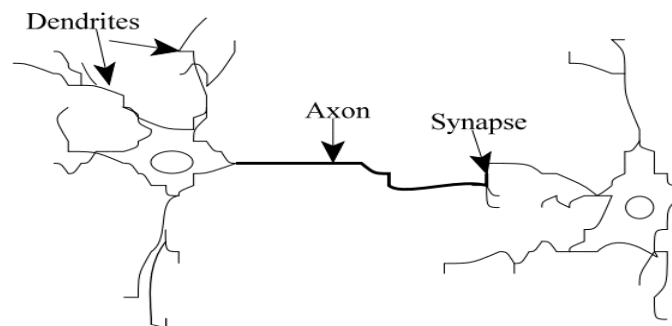


Fig. (1): Natural neurons

Neural networks can explore and analyze large and complex data, which is unusual in traditional models [13].

McCulloch and Pitts proposed the first neural model (1943). Artificial neural networks are mathematically modeled based on the functionality of biological neural networks [14]. The ANNs structure, as shown in Figure (2), consisted of:

- 1- Inputs (like synapses) represent the features multiplied by weights (strength of the respective signals).
- 2- Hidden neurons: each neuron receives signals from inputs after adding Bias and using an activation function to create an output signal.
- 3- Output: represent the summation of the signals created by the hidden neurons.

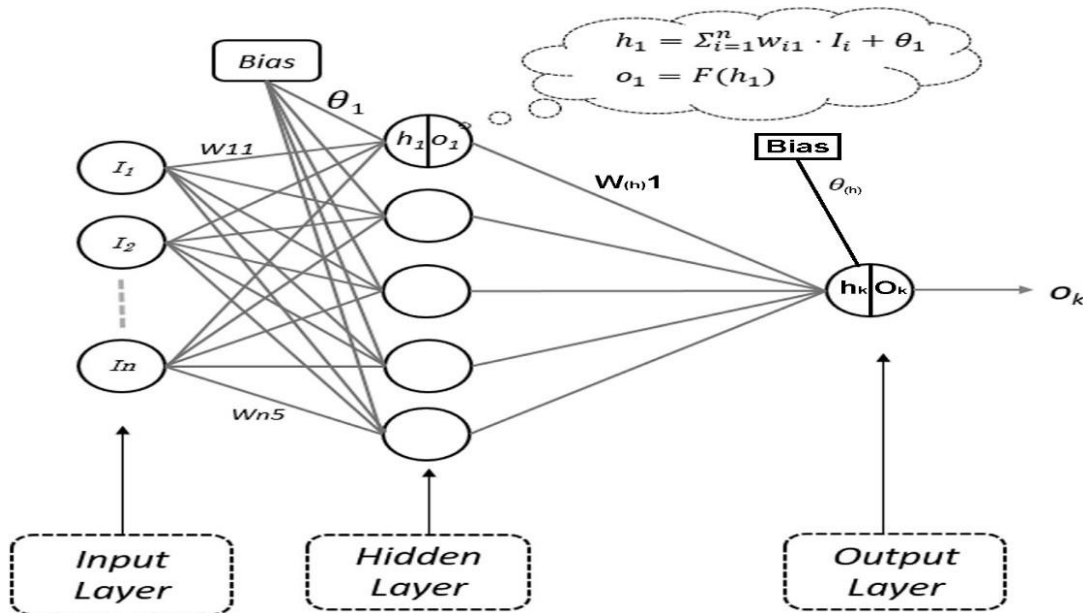


Fig. (2): Topology of a three-layer neural network [15]

Neural network topology defines the number of hidden layers and the number of neurons in each hidden layer used to connect the inputs to the outputs. The connections between every two nodes are called weights (w_{ij}), where i and j represent nodes in the input and output layer. Weights can be positive, strengthening the signal, or negative, weakening it [16]. In summary, the hidden nodes received the signal from the input nodes after multiplying by weight and adding Bias using the activation function to create the output signal as in (1) and (2).

$$h = \sum_{i=1}^n w_{ij} \cdot I_i + \theta_j \tag{1}$$

Where (h_j) represents the summation of the inputs (I_i) multiplied by weights(w_{ij}) added to Bias(θ_j), and (N) represents the total number of inputs.

$$O_j = F(h_j) \tag{2}$$

The O_j represents the nodes' output, and F is the activation function of h_j .

Repeat this process for each neuron in the first hidden layer. The outputs of the first hidden layer would be the inputs of the second hidden layer. This process continues until the output of the final hidden layer is found (O_k) as in (3) and (4). This process is called "feed-forward" when the absence of feedback from the final output to the previous layers.

$$h_k = O_j \times W_{j(h)} + \text{bias} \tag{3}$$

$$O_k = F(h_k) \tag{4}$$

The technique of how these weights are optimized is called the Backpropagation technique.

1.2 Backpropagation technique

The neural network training process aims to adjust the weights to obtain accurate outputs. The back-propagation technique updates the weights to reduce the error rate, representing the difference between the actual values of the output and the values calculated by the neural network. Back-propagation consists of two stages; the first stage is a feed-forward, which imposes initial values for weights, finds the outputs, and the error rate. The other step is to return to the network and adjust the weights to reduce the error rate to reach the lowest possible error.

Gradient Descent is used to calculate weight changes. The most common method is finding the slope (first derivative) to reach the minimum of the complex functions [17]. The back-propagation procedure is as follows:

- 1- Calculate the error using the output and target value for each neuron in the output layer using (5):

$$e = (O_k - d_j)^2 \tag{5}$$

Where (e) represents the error, (O_k) output value, (d_j) desired value of the model, and (k) the output neuron.

- 2- Calculate all neurons' total error (E) in the output layer. If there are more outputs, use (6):

$$E = \sum_{k=1}^n (O_k - d_j)^2 \tag{6}$$

- 3- Adjust the weights using the method of gradient descent using (7):

$$\Delta w_{ji} = -\eta \frac{\partial E}{\partial w_{ji}} \quad (7)$$

where Δw_{ji} represents the adjustment of each weight depending on η , which represents the learning rate multiplied by the derivative of E concerning w_{ji} ($\frac{\partial E}{\partial w_{ji}}$). The learning rate ranges from 0 to 1 and represents the rate of adjustment of weights and learning speed. The low value of the learning rate leads to a slow learning process, and thus slow access to the global minimum, and the high value leads to speed of learning, and thus the network is unstable and may be stuck in the local minimum, as shown in Figure (3). The derivative of E concerning to w_{ji} can be found as follows:

Determine how much the error depends on the output using (8) and (9):

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial O_j} \frac{\partial O_j}{\partial w_{ji}} \quad (8)$$

Where

$$\frac{\partial E}{\partial O_j} = 2 \times (O_j - d_j) \quad (9)$$

Determine how much the output depends on the activation, which in turn depends on the weights using (10):

$$\frac{\partial O_j}{\partial w_{ji}} = \frac{\partial O_j}{\partial h_j} \frac{\partial h_j}{\partial w_{ji}} = O_j(1 - O_j) \times I_i \quad (10)$$

Furthermore, we can see that (from (9) and (10)):

$$\frac{\partial E}{\partial w_{ji}} = \frac{\partial E}{\partial O_j} \frac{\partial O_j}{\partial w_{ji}} = 2 \times (O_j - d_j) O_j(1 - O_j) \times I_i \quad (11)$$

Thus, the adjustment to each weight will be (from (7) and (11)):

$$\Delta w_{ji} = -2\eta(O_j - d_j)O_j(1 - O_j) \times I_i \quad (12)$$

Finally, find the new weight using (13):

$$w_{ji}^n = w_{ji}^{n-1} + \Delta w_{ji} \quad (13)$$

- 4- After finding the new weights, the training process will go feed-forward, find the output and error, and adjust the weights again until the lowest error value is reached when the output almost matches the target.

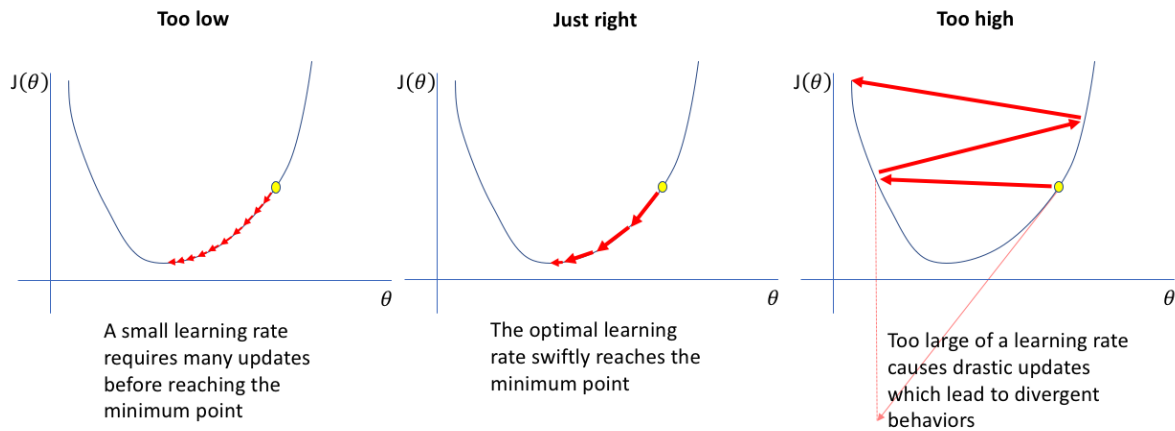


Fig. (3): Learning rate effect on training process [18]

1.3 Steps to train successful neural network model

1.3.1- Scale the data (Normalization)

1.3.2- Divide the data into training, testing, and validation.

1.3.3- Define the structure of the neural network (ANN topology), which includes (the No. of hidden layers, the No. of neurons in each layer, learning rate, activation or transformation function, solver, momentum, epoch, batch size, early stopping monitor, and patience) which explains as follows:

1.3.3.1. No. of hidden layers and neurons

There is no specific method to determine the number of hidden layers or neurons because it depends on the quality and complexity of the data used to train the model. Some programmers say that one hidden layer with specific neurons is enough to solve any problem, and there are some rules of thumb to determine the number of neurons in this single hidden layer:

- The number of hidden neurons should range between the number of neurons in the input and output layer.
- The number of the hidden neurons should be $2/3$ of the input neurons plus the number of the neurons in the output layer.

In complex problems, it must be taken into consideration that a small number of Neurons or hidden layers may lead to underfitting, which means that the model cannot train well and know the patterns. The excessive number of Neurons or hidden layers leads to overfitting, resulting in

a complex model with high training accuracy, but it fails to predict using any new data, as shown in Figure (4).

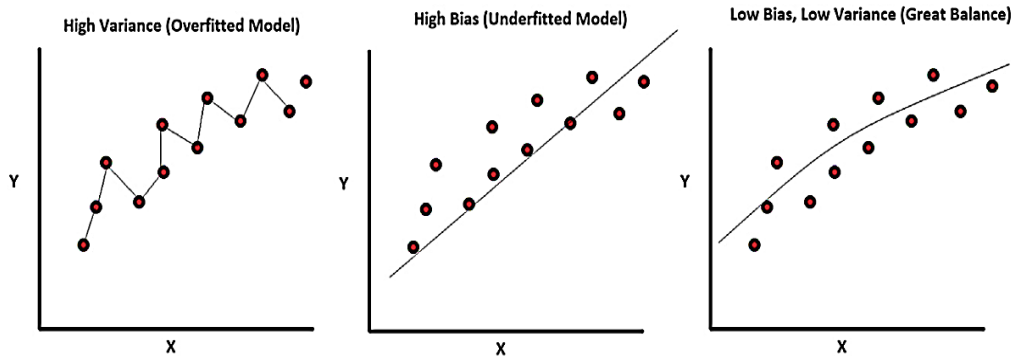


Fig. (4): Overfit, underfit, and good balance model [15]

1.3.3.2. Batch size: the number of samples per gradient update (number of inputs taken in each epoch).

1.3.3.3. Early stopping monitor and patience: early Stopping is used to monitor the decreasing error rate for training or testing to stop the training process when there is no improvement or increase in error rate [19]. It is imperative to monitor the Testing Process (val_loss) because the training process continues to evolve with the increase in the number of epochs, as shown in Figure (5). The training process will stop at 80 epochs when early Stopping monitor the testing process. The number of epochs with no improvement in error rate after which training will be stopped is called patience. The methodology will explain learning rate, activation or transformation function, solver, and momentum.

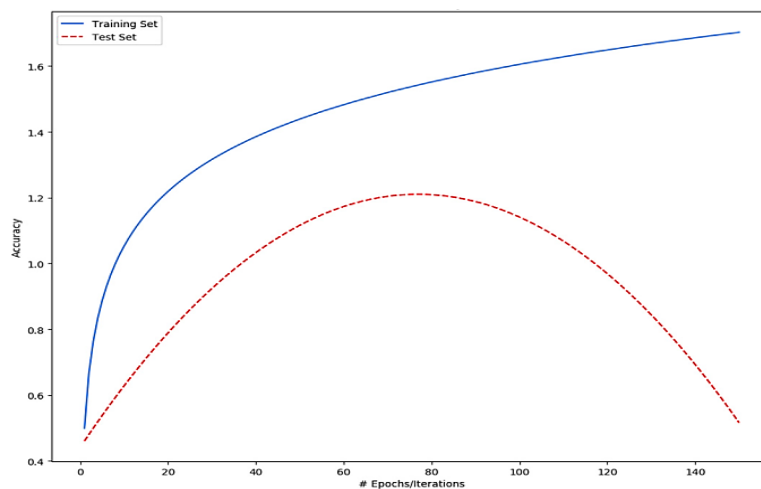


Fig. (5): Training and testing data set accuracy versus the number of iterations

- 1.3.4- Initialize weights and bias values (randomly).
- 1.3.5- Calculate the final output of the network (feed-forward) using Eqs.1,2,3 and 4.
- 1.3.6- Calculate the total error of the network using Eqs.5 and 6.
- 1.3.7- Adjust the weights (Back-propagation) using Eq. 13 (passing through Eq. 7 to Eq.12).
- 1.3.8- Repeat steps 5, 6, and 7 to minimize the total error as much as possible and stop the training.

2. Methodology

To predict lost circulation events in the Dammam formation: two artificial neural network models were developed based on drilling fluid properties, and drilling parameters for 75 wells drilled in the Rumaila oilfield passed through the Dammam formation, as shown in Table (1).

Table (1) Parameters used in modeling

No.	Parameter	Unit	Min.	Max.
1	Mud weight (MW)	gm/cc	1.05	1.09
2	Equivalent Circulating Density (ECD)	gm/cc	1.06	1.11
3	Yield Point (YP)	Ib/100ft ²	20	32
4	Plastic Viscosity (PV)	CP	6	18
5	Pump flow rate (Q)	L/STK	1760	3160
6	Rate Of Penetration (ROP)	m/hr	5	13
7	Revolutions Per Minute (RPM)	rpm	55	90
8	Stroke Per Minute (SPM)	spm	100	180
9	Total Flow Area (TFA)	in ²	0.547	6.643
10	Weight On Bit (WOB)	Ton	5	17
11	Loss Of Circulation (Loss)	m ³ /hr	No losses	91

The first model was developed using Tensor Flow with the Early Stopping Keras-Callback technique to stop the training process when the error rate increases or does not improve and lose the global minimum [19]. This technique helps to reach the global minimum without depending on specific *Epochs* (number of iterations which consisted of feed-forward, determining the error, and back-propagation) and get the best accuracy of the model. Early Stopping is very sensitive when using complex data in the training process when stopping at a local minimum instead of the global minimum. Momentum plays an important and sensitive role in the training process to help reach the global minimum. As it works to direct the training process toward the global minimum, bypassing the local minimum as shown in Figure (6). The momentum (μ) multiplied

by the previous rate of the weight adjustment to get a new rate this new rate helps to get the global minimum faster and pass local minimum as in (14).

$$\Delta_{ji}^n = \Delta_{wji} + \mu \cdot \Delta_{wji}^{n-1} \tag{14}$$

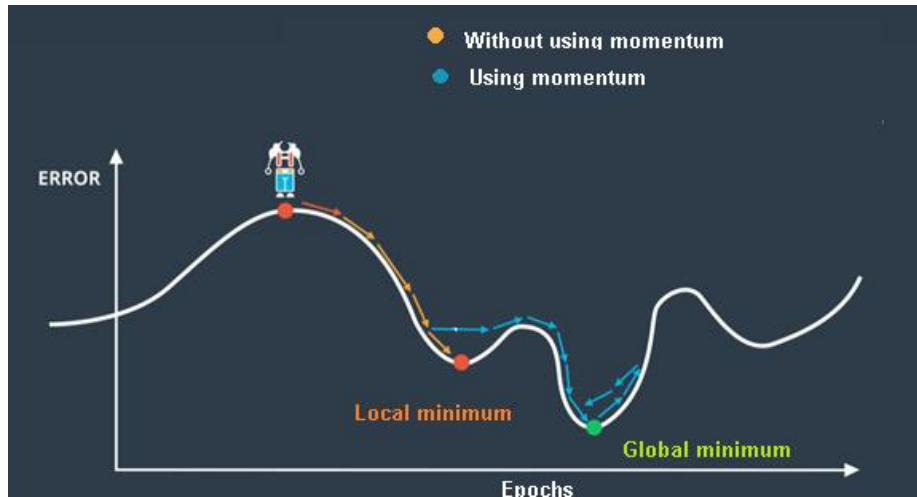


Fig. (6): Effect of momentum on the training process

The second model was developed using Tensor Flow with specific epochs and without early stopping to reach the global minimum. The best learning rate for the two models has been chosen using Learning Rate Scheduler [20], as shown in Figure (7).

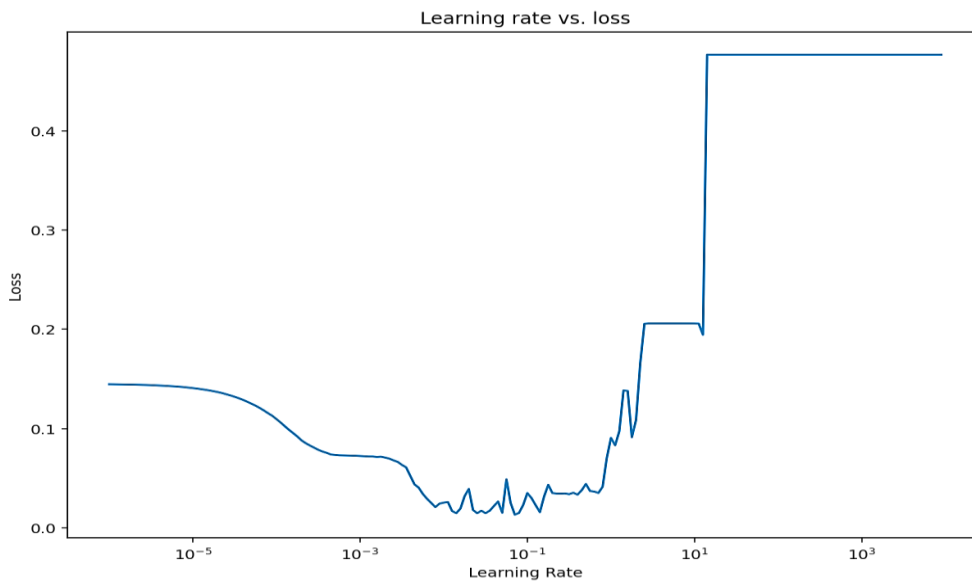


Fig. (7) Mean square error versus learning rate

The two models are summarized in Table (2) and Table (3).

Table (2) The best parameters of the two models

Model	No. of inputs	Early Stopping monitor	Patience	Learning rate	Training /testing ration	Epochs	Batch size	Output
1st model	10	Val_loss	30	0.0316	70/30	500	10	Lost circulation rate
2nd model	10	-	-	0.0316	70/30	500	10	Lost circulation rate

Table (3) The summary of the two models

Model	Input layer	1 st hidden layer		2 nd hidden layer		Output layer	
	Neurons	Neurons	Activation function	Neurons	Activation function	Neuron	Activation function
1st model	10	8	“ReLU”	4	“sigmoid”	1	“sigmoid”
2nd model	10	8	“ReLU”	4	“sigmoid”	1	“sigmoid”
	Total params:	129	Trainable params:	129	Non-trainable params:	0	

3. Results and Discussion

3.1. Model performance

After defining the best topology of each model using the Grid search technique [21], which operates a specific number of loops through which the hyperparameter values are replaced according to the values imposed by the user checking the accuracy of the model best hyperparameter can be selected), the training process for each model is shown in Figures (8) and (9).

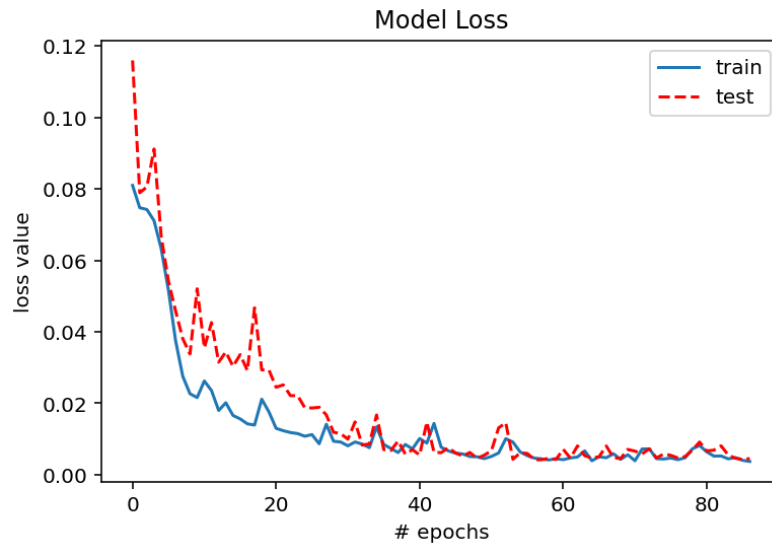


Fig. (8): Training and testing data set loss value versus epochs of the 1st model
Early stopping technique stop training at 87 epochs when the loss value of testing begins to increase.

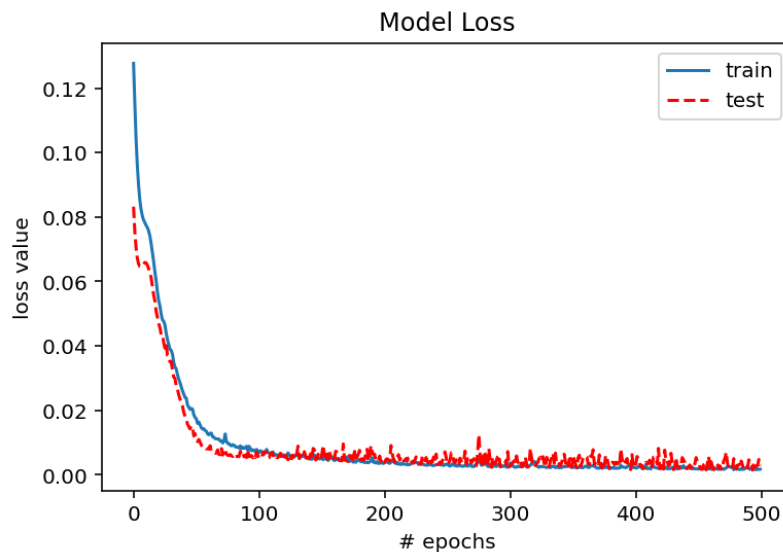


Fig. (9): Training and testing data set loss value versus epochs of the 2nd model

To evaluate the trained models using the R^2 metric for accuracy score and MAE, MSE, and RMSE metrics for error score, as shown in Table (4). After training the 1st model many times, the best accuracy was about 0.9485. That is because the complexity of the data makes the early stopping technique stop the training process at local minimum instead of global minimum and stuck their momentum. Patience plays a vital role in making early stopping technique stop the training process in the right place, which means the global minimum of the function.

Table (4) The performance of the two models

Model	Accuracy score (R ²)		Error score		
	Training	Testing	MAE	MSE	RMSE
1 st model	0.9485	0.9441	0.04921	0.00448	0.0669
2 nd model	0.9802	0.9567	0.03864	0.00279	0.05285

3.2. Implement the trained models

Implementing the trained model is the most essential step through which it determines whether the model is successful in predicting data that it has not seen before or fails to do so. Applying the data of 8 wells suffering from various types of lost circulation events in Dammam formation to predict these losses using the two models shows that the accuracy of implementing R² for the first and second models were 0.9302 and 0.9493, respectively. Figure (10) shows the actual and predicted losses value using the two models. The two models were able to predict the various type of losses ranging from seepage near-zero (m³/hr) to total losses of about 87(m³/hr) with high accuracy.

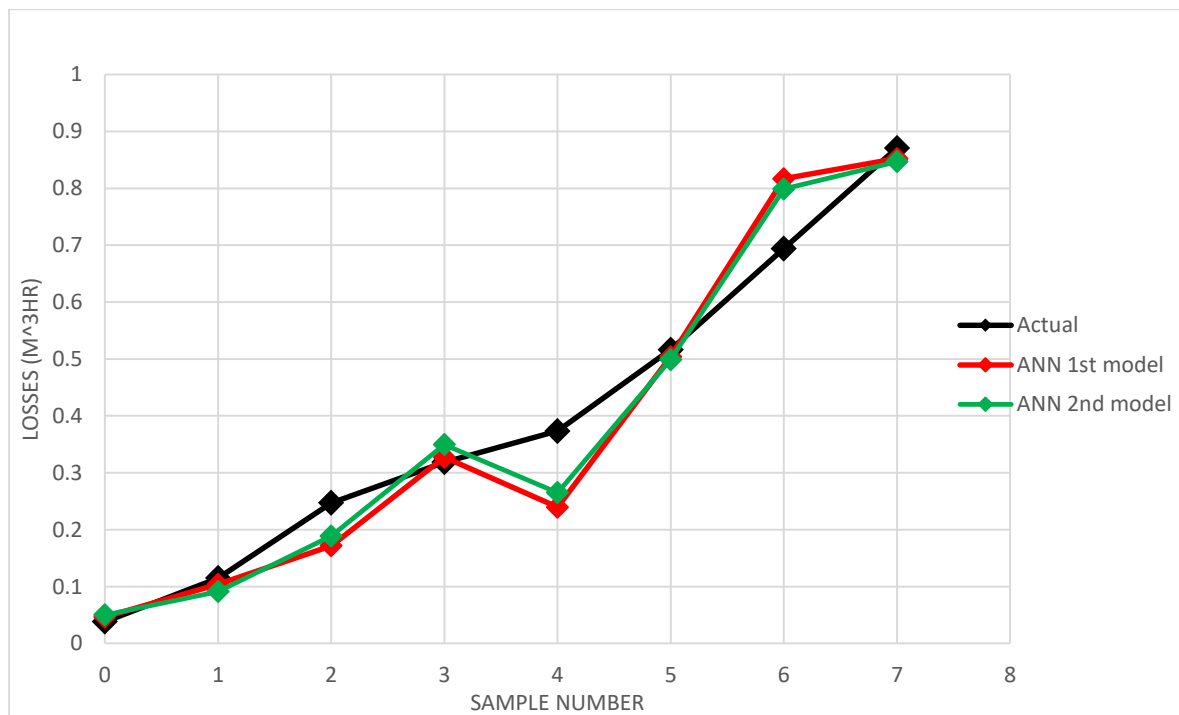


Fig. (10): Comparison of the estimated values of each model and the real losses (the values are under normalization)

4. Conclusion

- Drilling fluid loss is a complex problem that is difficult to predict using simple and traditional methods. Neural networks represent a modern and accurate technology for solving complex problems such as drilling fluid loss.
- Hyperparameters of the two models were selected using the Grid search technique.
- The early stop technique allows the intelligent model to be trained with high accuracy and in a short period without dependence on the number of epochs.
- The results were perfect, the built models were applied, and they accurately predicted all types of losses.
- These models have been prepared to predict the loss of drilling fluid that occurs in the Dammam formation- Rumaila oil field only, and they cannot be used in another field, even if the field is very close.

Abbreviation

AI	Artificial Intelligence
ANN	Artificial Neural Network
ECD	Equivalent Circulating Density
LCM	Lost Circulation Materials
MAE	Mean Absolute Error
MSE	Mean Square Error
MW	Mud weight
NPT	Non-Productive Time
PV	Plastic Viscosity
ReLU	Rectified Linear Unit
RMSE	Root Mean Square Error
ROP	Rate Of Penetration
RPM	Revolutions Per Minute
SPM	Stroke Per Minute
TFA	Total Flow Area
WOB	Weight On Bit
YP	Yield Point

Nomenclature

η	Learning rate
μ	Momentum
Q	Pump flow rate
R^2	Square Linear Correlation Co-Efficient Value

References

- [1] Giin-Fa Fuh, Nobuo Morita, P. A. Boyd, and S. J. McGoffin, "A New Approach to Preventing Lost Circulation While Drilling", Paper presented at the SPE Annual Technical Conference and Exhibition, Washington, D.C., October 1992. doi: <https://doi.org/10.2118/24599-MS>
- [2] R. Caenn, H. C. H. Darley, and G. R. Gray, "Composite and properties of Drilling and Completion Fluids", [Book]. – six edition, 2011. <https://doi.org/10.1016/C2015-0-04159-4>
- [3] O. Razavi, A. K. Vajargah, E. van Oort, and M. Aldin, "Optimum particle size distribution design for lost circulation control and wellbore strengthening", Journal of Natural Gas Science and Engineering, vol. 35, pp. 836–850, 2016. <http://dx.doi.org/10.1016/j.jngse.2016.08.038>
- [4] C. Ezeakacha, S. Salehi, "Experimental and statistical investigation of Drilling Fluids Loss in porous media– part 1", Journal of Natural Gas Science and Engineering, vol. 51, pp. 104-115, 2018. <http://dx.doi.org/10.1016/j.jngse.2017.12.024>.
- [5] M. Alsaba, M. Al Dushaishi, R. Nygaard, O. Nes, and A. Saasen, "Updated criterion to select particle size distribution of lost circulation materials for an effective fracture sealing", Journal of Petroleum Science and Engineering, vol. 149, pp. 641–648, 2017. <https://doi.org/10.1016/j.petrol.2016.10.027>
- [6] U. Arshad, B. Jain, M. Ramzan, W. Alward, L. Diaz, I. Hasan, A. Aliyev, and C. Riji, "Engineered Solution to Reduce the Impact of Lost Circulation During Drilling and Cementing in Rumaila Field, Iraq", Paper presented at the International Petroleum Technology Conference, Doha, Qatar, December 2015. doi: <https://doi.org/10.2523/IPTC-18245-MS>
- [7] J. U. Messenger, J. S. McNiel, "Lost Circulation Corrective: Time-Setting Clay Cement", Journal of Petroleum Technology, 4(3), pp. 59-64, 1952.
- [8] A. Al-Hameedi, H. H. Alkinani, S. Dunn-Norman, R. E. Flori, and S. A. Hilgedick, "Real-time lost circulation estimation and mitigation", Egyptian journal of petroleum, vol. 27, no. 4, pp.1227-1234, 2018. doi: <https://doi.org/10.1016/j.ejpe.2018.05.006>.
- [9] A. R. Moazzeni, M. Nabaei, and S. Ghadami Jegarluei, "Prediction of Lost Circulation Using Virtual Intelligence in One of Iranian Oilfields", In: Proceeding of the SPE Nigeria Annual International Conference and Exhibition, Tinapa – Calabar, Nigeria, 31 July–7 August 2010. doi: <https://doi.org/10.2118/136992-MS>.

- [10] H. Toreifi, H. Rostami and A. K. manshad, "New Method For Prediction And Solving The Problem of Drilling Fluid Loss Using Modular Neural Network and Particle Swarm Optimization Algorithm", Journal of Petroleum Exploration and Production Technology, vol. 4, no. 4, pp. 371–379, 2014. doi: <http://dx.doi.org/10.1007/s13202-014-0102-5>.
- [11] A. K. Abbas, U. Alameedy, M. Alsaba, and S. Rushdi, "Wellbore Trajectory Optimization Using Rate of Penetration and Wellbore Stability Analysis", In: SPE International Heavy Oil Conference and Exhibition, Kuwait City, Kuwait, 10-12 December 2018. doi: <http://dx.doi.org/10.2118/193755-ms>.
- [12] S. Mohaghegh, D. McVey, K. Aminian; S. Ameri, "Predicting Well-Stimulation Results in a Gas-Storage Field in the Absence of Reservoir Data with Neural Networks", SPE-31159-PA, SPE Reservoir Engineering, vol. 11, no. 04, pp. 268-272, 1995.
- [13] S. Mohaghegh, "Virtual-intelligence applications in petroleum engineering: Part 1d artificial neural networks", Journal of Petroleum Technology, vol. 52, no. 09, pp. 64-73, 2000. <https://doi.org/10.2118/58046-JPT>
- [14] W. S. McCulloch, W. Pitts, "A logical calculus of the ideas immanent in nervous activity", Bulletin of Mathematical Biophysics 5, pp. 115–133, 1943. <https://doi.org/10.1007/BF02478259>
- [15] H. Belyadi, A. Haghghat, "Machine Learning Guide for Oil and Gas Using Python: A Step-by-Step Breakdown with Data, Algorithms, Codes, and Applications", Gulf Professional Publishing, an imprint of Elsevier. 2021. <https://doi.org/10.1016/C2019-0-03617-5>
- [16] D. Kriesel, "A brief introduction to neural networks", 2005. www.dkriesel.com.
- [17] L. Fausett, "Fundamentals of neural networks: Architectures, algorithms and applications", 1993.
- [18] J. Jordan, "Setting the learning rate of your neural network", Retrieved 1 MAR 2018, from <https://www.jeremyjordan.me/nn-learning-rate/>
- [19] Keras Team (n.d.). Keras documentation: Early Stopping. Retrieved 21 May 2022, from https://keras.io/api/callbacks/early_stopping/
- [20] Keras Team (n.d.). Keras documentation: Learning RateScheduler. Retrieved 21 May 2022, from https://keras.io/api/callbacks/learning_rate_scheduler/
- [21] Scikit-learn (n.d.). Sklearn model selection Grid search CV. Retrieved July 17, 2020, from https://scikitlearn.org/sklearn.model_selection.GridSearchCV