The International Journal of Engineering and Information Technology

journal homepage:www.ijeit.misuratau.edu.ly

A taxonomy of Collective Machine Learning Models applicable to well logging predictive anticipation

Elshebani, M.
Libyan Academy, Tripoli, Libya.
Mohamed.elshbani@academy.edu.ly.

Alnour, S.
Libyan Academy, Tripoli, Libya.
220100899@academy.edu.ly.

Abstract—**Traditional methods of measuring well logs are expensive, error-prone and time-consuming, which has led to the development of machine learning models that can predict well logging based on well-log data. This study aims to determine the most effective machine learning models for predicting of well logging based on available well-log data. The study covers a detailed explanation of the data-gathering and pre-processing techniques used.**

Features were used in the models, namely gamma ray (GR), bulk density (RHOB), neutron porosity (NPHI), resistivity (RT), spontaneous potential (SP), trained and evaluated based on their performance, namely linear regression, support vector machine SVM, Neural Network NN and decision Trees DT models. The models were evaluated based on their Mean Squared Error, R squared, Mean Absolute Error and RMSE values. Our results showed that the Decision Trees (DT) for MSE value of 10.86, achieving a Root Mean Squared Error (RMSE) value of 3.29, MAE value of 2.225 and R square value of 0.59. These findings suggest that machine learning models can be a powerful tool for predicting of best training from well-log data, in particular, holds great promise for future modelling efforts in this area.

Index Terms— **Machin learning, Well logging, Training Data, Predication, RMSE, MSE, MAE, R square.**

I. INTRODUCTION

Many real-world situations might benefit from the use of machine learning (ML) approaches due to recent advancements in ML and rising computational capacity. With its abundance of data, the petroleum sector is well-positioned to take advantage of these strategies and provide major benefits. The measurements of physical attributes that are obtained during the drilling of exploratory boreholes and documented in well logs sequential recordings of features collected at regular depth increments usually make up the accessible data.

Received 30 Apr , 2024; revised 11 May, 2024; accepted 15 Mar 2024.
Available online 08 Aug, 2024

The features of the surrounding geology are inferred from the well logs using a variety of modeling methods. In the end, these help with commercial decision-making about the well's further development or the development of whole oilfields, depending on the hydrocarbon content estimate.

Well log interpretation is a time-consuming and costly process that transforms raw data into information that is useful to the commercial world. It calls for a large amount of human labor as well as a high level of skill and experience. Inaccurate hydrocarbon content estimation also has a big financial impact as it can lead to lost chances or high drilling expenses for a producing well with poor hydrocarbon return.

The well logs, which are inherently noisy and imperfect, include raw measurements taken with a variety of instruments. Petro-physicists were assigned the duty of correcting incorrect readings and guessing missing data in order to "condition," or clean, the well logs. Only once this has been completed, they can proceed with the interpretation of the rock properties.

For a Petro-physical assessment of subsurface hydrocarbon-bearing formations, a number of logs are needed, including resistivity (RACEHM, RACELM, RD, RM, RPCEM, RPCEM, RT), density (RHOB), self-potential (SP), sonic (DT), gamma ray (GR), neutron (NPHI), resistivity (CALI), and so on. These fundamental logs may be used to estimate reservoir thickness with fluid, water saturation, clay volume, grain size, porosity, and borehole diameter.

In this study, we determine the relationships among several physical characteristics measured within a borehole using supervised machine learning approaches. One of these attributes is then used as a dependent or target variable, while the other attributes are used as independent or input variables. Each well log record corresponds to a single observation, and we use a subset of these observations to train the models while keeping an evaluation-ready disjoint set. We only use data if the desired attribute has been recorded for training and evaluation purposes, utilizing the measured value as the ground truth. Petro-physicists will no longer need to label the data in advance, and our theory is that this will lead to more consistent data values and prevent human intervention.

A. Previous studies

Pan, Tan, and Hu (2009) developed a mathematical model and numerical method for analyzing spontaneous potential logs in heterogeneous formations [7]. Gąsior and Przelaskowska (2014) estimated thermal conductivity from core and well log data. Their work provided methods for better thermal conductivity estimation, crucial for various geological and engineering applications [12]. Khisamutdinov and Banzarov (2015) addressed mathematical modeling issues related to pulsed neutron-gamma logging [10]. Zhang, Chen, and Meng (2018) utilized Recurrent Neural Networks (RNNs) to generate synthetic well logs [8]. Kerimova (2019) applied mathematical techniques and integrated well logging data to classify oil-gas bearing targets based on their saturation features and facies composition in her study. Her research contributed to more accurate classification of hydrocarbon-bearing targets [11]. Ahmadi and Chen (2019), identified the most accurate models for these predictions, contributing to better resource estimation in oil reservoirs [3].
Nordloh, Roubícková, and Brown (2020) discussed the application of machine learning techniques in gas and oil exploration and, this study provided a broad overview of how machine learning can be utilized in exploration activities, emphasizing its potential in improving exploration efficiency [4].
Joshi et al. (2021) in their study applied various machine learning techniques to predict sonic logs and correlate lithology using geophysical well log data [1]. Talebkeikhah, Sadeghtabaghi, and Shabani (2021), this research compared different machine learning models for predicting permeability in hydrocarbon reservoirs using well log data, showcasing the relative performance of various approaches [2]. Ullah et al. (2023) presented a multidisciplinary approach to facies evaluation at a regional level in their study. This study combined well log analysis, machine learning, and statistical methods to enhance regional facies evaluation [6]. Rashidi et al. (2023) discussed common statistical concepts in supervised machine learning, including those in well logging [13]. Rahmati, Zargar, and Tanha (2024) in their comparative study, their research identified the most effective methods for this prediction task [5].

THE BASIC CONCEPTS

The basic concepts related to the paper that will help us to understand and comprehend the paper very well.

A. Machine Learning

Using a collection of existing training data, supervised machine learning (ML) automatically determines a functional link between input and output variables. Each input comprises an ordered vector of values, referred to as independent variables or features, that characterize different aspects of the issue. Through the evaluation of the learnt function across the input vector, the output values also referred to as dependent or target variables are predicted [5].
Regression is used for goals with a continuous, possibly infinite domain, whereas classification is used for targets where the target values constitute a (often small) finite set. This distinction is based on the domain of the target variables.
We used neural networks, Support Vector Machines and gradient tree boosting as two regression models to approximate the missing data in the logs. Both models are essentially different, even though they both take a vector of characteristics as input and output a goal value.
A neural network (NN) is made up of many layers, each of which has a variable number of neurons.
A neural network with one hidden layer is shown in Figure 1, and the network's exact design is determined by the issue it is meant to solve. Each neuron shown in Figure 2, the first input layer represents a single value from the input vector. In the training phase, these weights are learnt. The values of neurons in subsequent layers are computed as a weighted sum of their predecessors modified by an activation function.
The last layer, called the output layer, shows the inferred value, which is derived from the vector of features supplied to the network, the activation functions (which specify a node's output based on its input), and the weights. At a neuron's output, the activation function serves as a decision-making body [8]. Based on the activation function, the neuron learns either linear or non-linear decision limits. Additionally, because of the cascading effect, it exerts a leveling impact on neuron output, preventing neurons' output from growing excessively big after several layers [6].
Neural networks can detect complex patterns in extremely non-linear data sets, such as well logs, since the activation function does not have to be linear to its parameters.

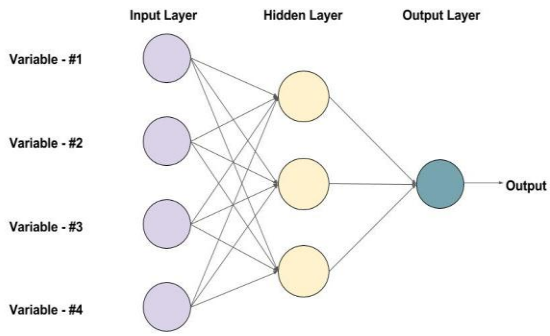


Figure 1: Neural Network with one hidden layer (3 neurons)

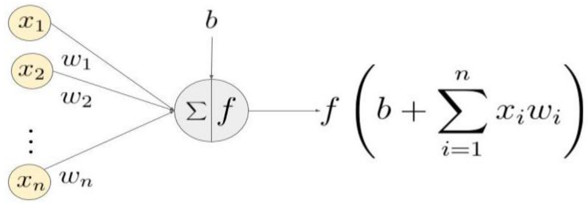


Figure 2: a neuron showing the input (x1-xn), their corresponding weights(w1-wn), a bias (b) and the activation function (f) applied to the weighted sum of the input.

Gradient Boosting is an ensemble learning technique that builds a powerful prediction model by combining several weak prediction models, often decision trees. Regression, classification, and ranking issues are just a few of the fields in which it has been effectively employed in machine learning. They frequently perform better than other conventional models and are strong algorithms for handling intricate forecasting jobs.

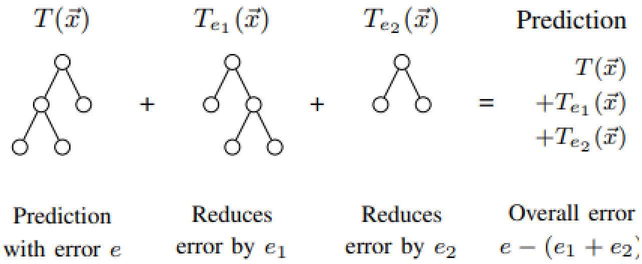


Figure3: The first tree (T) is a decision tree providing a rough estimation of the value to be predicted. The second tree (Te1) predicts and corrects the error of T, leading to a reduced overall error [4].

The accuracy of these regression decision trees is severely constrained since they can only reliably forecast the values that are represented by their leaves. Gradient Tree Boosting gets around this restriction by building an ensemble of trees, as shown in Figure 3, where each tree that follows after the first corrects the error of the one before it, improving the prediction until the error converges or the model begins to overfit.

B. Well Logging

Nowadays, advanced and modern techniques of Well logging are considered extremely useful for subsurface formation evaluation and fluid properties predictions. Various logs required for a petrophysical evaluation of subsurface hydrocarbon-bearing formations are Caliper (DCAL, MCAL), Delta Rho (DROR), Gamma Ray (GR), Neutron Porosity (RHOC, CNLS), Density (RHOB), Spontaneous Potential (SP), Resistivity (RLL3, RILD, MI, MN, RxoRT), etc.

S.N	LOG TYPE	DESCRIPTION
1	Caliper Log	• Measure the diameter along the length of borehole • Contributory information for lithological assessment • Indication and thickness of permeable zones
2	Gamma Ray Log	• Gamma ray responses of formation by detecting radioactivity • Design of lithological profile and lithological correlation between wells • Differentiate between reservoir and non-reservoir rock • Estimate shale content and applicable for depth matching
3	Resistivity Log	• Determination of hydrocarbon versus water bearing zones • Determination of resistivity, porosity and water saturation • Measurement of resistivity of flushed, shallow and deep zones
4	Bulk Density Log	• Depends on gamma ray scattering and photoelectric absorption • Measurement of Bulk Density of formation • Determination of porosity and differentiate between oil bearing and gas bearing zones
5	Neutron Porosity Log	• Measurement of hydrogen concentration in the formation • Measurement of total porosity • Shale increases measured actual porosity and gas reduces measured actual formation porosity

Figure 4: Details of the primary type of well logs used in the Exploration and Production (E&P) industry [1].

II. THEORETICAL BACKGROUND

Basic ideas and techniques form the theoretical basis for an extensive comparative analysis of several machine learning models in well log prediction and evaluation [4]. An important component of subsurface exploration is well logging, which procedures a variety of characteristics of rocks and fluids, as well as gamma-ray, resistivity, porosity, and acoustic logs [9]. By analyzing the physical and chemical characteristics of the materials that make up the Earth's subsurface, the field of petrophysics is important to the interpretation of well log data. Regression models such as support vector regression, decision tree regression, and linear regression are frequently used in supervised learning to make predictions, whereas classification models are used to handle discrete outcomes such as petrology classes [13]. Log transformations and depth-related features are two feature engineering strategies that improve the model's capacity to represent geological patterns.

Normalization, scaling, and filling in missing data are examples of data preprocessing procedures that guarantee the constancy of machine learning algorithms, various complexities in geological datasets in [4] are accommodated by model selection procedures such as decision trees, support vector machines, and neural networks. While classification metrics calculate accuracy, precision, recall, and F1-score for classification tasks, evaluation metrics like Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared assistance analyze prediction accuracy [3]. By dividing the dataset into many folds, cross-validation techniques such as K-Fold Cross-Validation ensure robust model evaluation [12-13]. The models are further refined through ensemble approaches, hyperparameter tuning, and concerns for interpretability and integration of domain knowledge. Specialized strategies are needed to address imbalanced data, particularly in situations where certain lithologies are less common. investigation.

III. METHODOLIGY AND INTERPERTAION

The paper focused on analysis of well log data using various ML approaches forms the basis of the current work. The utilized data sets were gathered from the Kaggle website's open-sourced inventory. Excel format is used to see the log data. The header includes the fundamental data, such as the name of the well and the depth measurements of the borehole's beginning and finish. CALI, GR, RHOB, DT, NPHI, and RT are the different logs that were utilized to build the inventory. A range of methodologies are employed to identify any absent numbers, comprehend the distribution of the data, and present the accessible data visually.

This study aims to following key challenges:

The aim to Evaluate and compare the performance of various machine learning models, such as linear regression, decision trees, support vector machines, neural networks, and ensemble methods, in predicting well log properties.

The first, Data Collection by collect a various and dataset of well logs containing geological formations and confirm the addition of relevant well log properties such as Gamma-ray, Porosity, Resistivity, and Permeability and Data Preprocessing or cleaning by clean the dataset by handling missing values, and any discrepancies.

The next, Model Selection by take a set of various machine learning models for comparison, including linear regression, decision trees (DT), support vector machines (SVM), neural networks (NN) and apply Model Training by train each selected machine learning model using the training set, and Optimize model hyperparameters, the last, Model Evaluation by assess the models on the testing set using suitable metrics (e.g., mean squared error, R-squared) to measure predictive performance.

IV. MATHMATICAL MODELLING

The literature on numerical techniques and error analysis for the SP log issue, such as the numerical mode matching (NMM) approach and finite element method (FEM), is extensive since SP log is a crucial technology in petroleum extraction [10]. The NMM approach is significantly simpler than the FEM method [7].

As demonstrated in described [7] by equation

$$R_e = \begin{cases} R_m & in\ \Omega_m \triangleq \Omega_1 \\ R_s & in\ \Omega_s \triangleq \Omega_2 \\ R_f(r) = \left(\dfrac{1}{R_t} + \dfrac{\frac{1}{R_x} - \frac{1}{R_t}}{1 + \left(\frac{r - r_o}{w}\right)^N} \right)^{-1} & in\ \Omega_t \triangleq \Omega_3 \end{cases} \tag{1}$$

Where:
r_o is the radius of the well,
w is the radius where the half of the transition from R_x to R_t occurs, and
N is directly related to the slope of the transition.
Ω_m is the well bore filled with mud, Ω_s is the enclosing rock, and Ω_t is the objective layer which contains the invasion zones.

It is difficult to generalize in order to solve the SP log issue for the scenario in which the formation resistivity R relies on r as shown in equation (1). But FEM is a more precise approach than NMM.

A. Difference schemes in subdomains

Since the resistivity in each subdomain Ω_i $(i = 1,2,3)$ depends only on r, let $R_i = R_e\ (r_i)$, we use the standard 5-point scheme [7]:

$$r_i = i\Delta r, \quad z_j = j\Delta z \quad i = 0,1,\ldots M, \quad j = 0,1,\ldots N$$

$$\frac{\beta_{i+\frac{1}{2}}\mathcal{U}_{i+1,j} - \left(\beta_{i+\frac{1}{2}} + \beta_{i-\frac{1}{2}}\right)\mathcal{U}_{i,j} + \beta_{i-\frac{1}{2}}\mathcal{U}_{i-1,j}}{\Delta r^2} + \beta_i \frac{\mathcal{U}_{i,j+1} - 2\mathcal{U}_{i,j} + \mathcal{U}_{i,j-1}}{\Delta z^2} = 0 \tag{2}$$

Where: $\boldsymbol{\beta}_i = \frac{\boldsymbol{r_i}}{\boldsymbol{R_i}}$

$\boldsymbol{\beta}_{i\pm\frac{1}{2}} = \frac{\boldsymbol{\beta_i + \beta_{i+1}}}{\boldsymbol{2}}$

B. Difference schemes on interfaces

Just the difference scheme on the horizontal interface, $z = H = j_o\Delta z$, needs to be taken into account. For the vertical interface, $r = r_o$, a similar process may be followed.

Let the resistivities above and below the interface $z = j_o\Delta z$ be denoted by R^+ and R^-, respectively. Utilizing the close-to-interface local Taylor expansions, we arrive at [7]

$$\mathcal{U}_{i,j_o+1} - \mathcal{U}_{i,j_o{}^+} = \Delta z\left(\frac{\partial \mathcal{U}}{\partial x}\right)_{i,j_o{}^+} + \frac{\Delta z^2}{2}\left(\frac{\partial^2 \mathcal{U}}{\partial z^2}\right)_{i,j_o{}^+} + O(\Delta z^3) \tag{3}$$

$$\mathcal{U}_{i,j_o-1} - \mathcal{U}_{i,j_o{}^-} = -\Delta z\left(\frac{\partial \mathcal{U}}{\partial x}\right)_{i,j_o{}^-} + \frac{\Delta z^2}{2}\left(\frac{\partial^2 \mathcal{U}}{\partial z^2}\right)_{i,j_o{}^-} + O(\Delta z^3) \tag{4}$$

From (3) and (4) it is obvious that

$$\begin{aligned} &\frac{1}{\Delta z^2}\left(\beta_i^+(\mathcal{U}_{i,j_o+1} - \mathcal{U}_{i,j_o{}^+}) + \beta_i^-(\mathcal{U}_{i,j_o-1} - \mathcal{U}_{i,j_o{}^-})\right) \\ &= \frac{1}{\Delta z}\left[\beta\frac{\partial \mathcal{U}}{\partial x}\right] + \frac{1}{2}(\beta_i^+\left(\frac{\partial^2 \mathcal{U}}{\partial z^2}\right)_{i,j_o{}^+} + \beta_i^-\left(\frac{\partial^2 \mathcal{U}}{\partial z^2}\right)_{i,j_o{}^-} + O(\Delta z) \\ &= -\tfrac{1}{2}\left(\tfrac{\partial}{\partial r}(\beta\tfrac{\partial \mathcal{U}}{\partial r})_{i,j^+} + \tfrac{\partial}{\partial r}(\beta\tfrac{\partial \mathcal{U}}{\partial r})_{i,j^-}\right) + \mathrm{O}(\Delta z) \\ &= -\frac{1}{2}\left(\frac{\beta^+_{i+\frac{1}{2}}\mathcal{U}_{i+1,j_o^+} - \left(\beta^+_{i+\frac{1}{2}} + \beta^+_{i-\frac{1}{2}}\right)\mathcal{U}_{i,j_o^-} + \beta^-_{i-\frac{1}{2}}\mathcal{U}_{i-1,j_o^-}}{\Delta r^2} \right. \\ &\left. + \frac{\beta^-_{i+\frac{1}{2}}\mathcal{U}_{i+1,j_o^+} - \left(\beta^-_{i+\frac{1}{2}} + \beta^-_{i-\frac{1}{2}}\right)\mathcal{U}_{i,j_o^-} + \beta^-_{i-\frac{1}{2}}\mathcal{U}_{i-1,j_o^-}}{\Delta r^2}\right) \\ &+ O(\Delta r + \Delta z) \end{aligned} \tag{5}$$

Thus, the resulting coefficient matrix of the finite difference method is symmetric and positive definite as shown in equation (5). The matrix can be inverted using common acceleration techniques like multigrid and preconditioned conjugate gradient (PCG) algorithms.

V. METRICS FOR MODEL EVALUATION

Define evaluation measures according to the type of task being predicted. Metrics like Mean Squared Error (MSE), Root Mean Squared Error (RMSE), R Square (R^2) and Mean Absolute Error (MAE) are frequently employed for regression jobs [13].

TABEL 1. DESCRIBE OF MSE, R SQUARE, RMSE AND MAE

Statistic	Description
RMSE	Root mean squared error. The RMSE is always positive and its units match the units of your response.
R-Squared	Coefficient of determination. R-squared is always smaller than 1 and usually larger than 0. It compares the trained model with the model where the response is constant and equals the mean of the training response. If your model is worse than this constant model, then R-Squared is negative.
MSE	Mean squared error. The MSE is the square of the RMSE.
MAE	Mean absolute error. The MAE is always positive and similar to the RMSE, but less sensitive to outliers.

A. Data Features

These data features collectively contribute to the interpretation of subsurface geological conditions and are vital for reservoir characterization in the oil and gas industry [11]. Machine learning models can be trained using these features to predict and evaluate well logs, facilitating more efficient and accurate decision-making in exploration and drilling operations [9].

DEPTH .FT :1 Depth
DCAL .IN :2 Caliper(Caliper – Density)
MCAL .IN :3 Caliper(Caliper – Sonic)
DROR.G/CC :4 Drho (Delta Rho)
MN . OHMM :5 Res(Resistivity Wide arry)
GR .GAPI :6 Gamma Ray (Gamma Ray)
RxoRt .OHMM :7 Res(Ratio of shallow and deep resistivity in well)
RILD .OHMM :8 Deep Res(Deep Induction Standard Processed Resistivity)
MI .OHMM :9 Med Res(Medium Induction Standard Processed Resistivity)
RLL3 .OHMM :10 Shal Res(Latero-Log 3)
CNLS.V/V :11 Neutron(Neutron Porosity)
RHOC.V/V :12 Neutron(Neutron Porosity)
RHOB .G/CC :13 Density(Bulk Density)
SP .MV :14 SP(Spontaneous Potential)

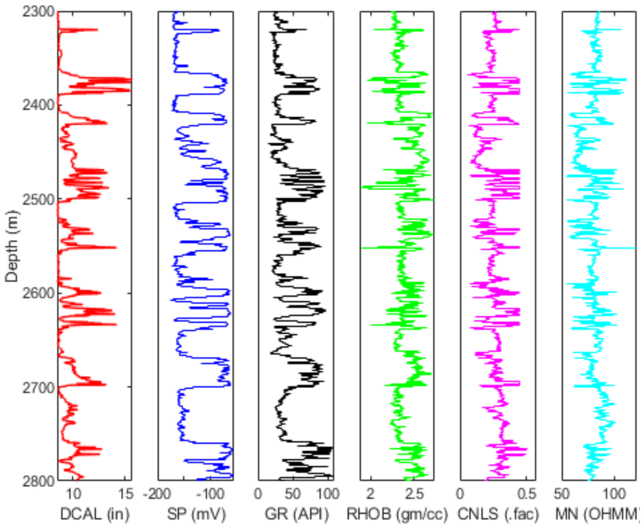


Figure5: display some of data features

VI. RESULTS AND DISCUTION

The regression leaner and classification leaner train models to predict data. we used training to search for the best regression model type, including linear regression models, regression trees, Gaussian process regression models, support vector machines, ensembles of regression trees, and neural network regression models.

The regression learner performs hyperparameter tuning by using Bayesian optimization. The goal of Bayesian optimization, and optimization in general, is to find a point that minimizes an objective function a point is a set of hyperparameter values, and the objective function is the loss function.

Machine learning models can be trained using features to predict and evaluate well logs, facilitating more efficient and accurate decision-making in exploration and drilling operations by using MATLAB R2022a.

TABEL 2. DESCRIBE THE MEASUREMENTS OF MSE, R SQUARE, RMSE AND MAE

	Regression Tree	Linear Regression	Wide Neural Network
MSE	11.2	53.6	12.3
R SQUARE	0.91	0.59	0.91
RMSE	3.34	7.32	3.59
MAE	2.280	5.661	2.500

A. *The Regression Leaner*

1- Regression Tree (Model 1)

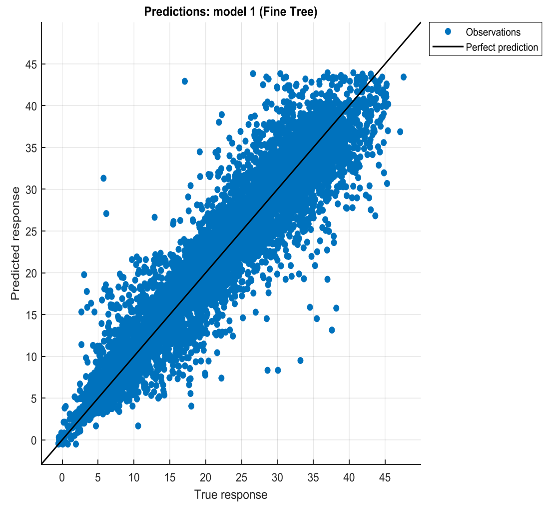


Figure 6: prediction of data.

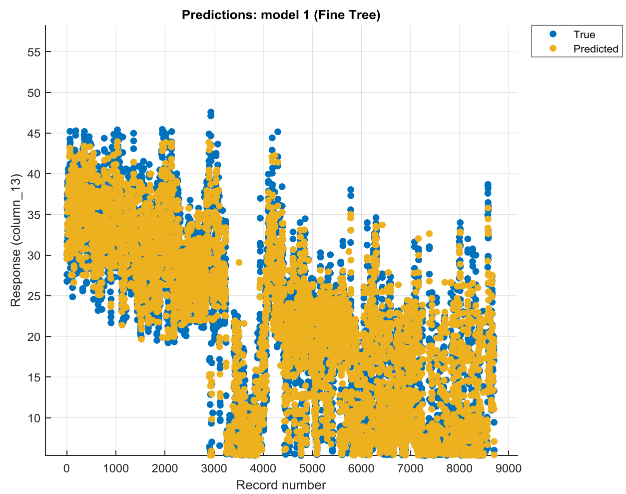
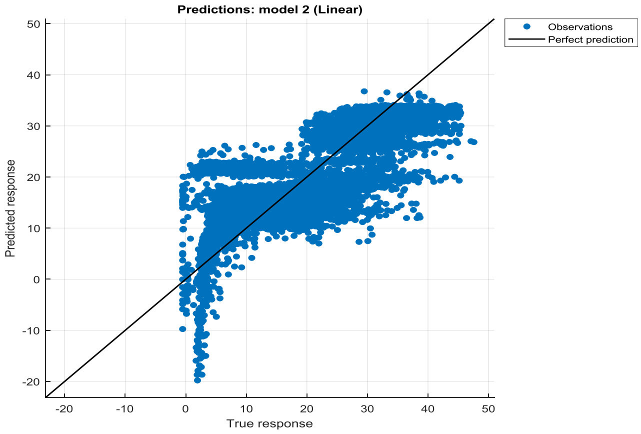


Figure7: true and predicted data

2 - Liner Regression(Model 2)



prediction of data :Figure 8.

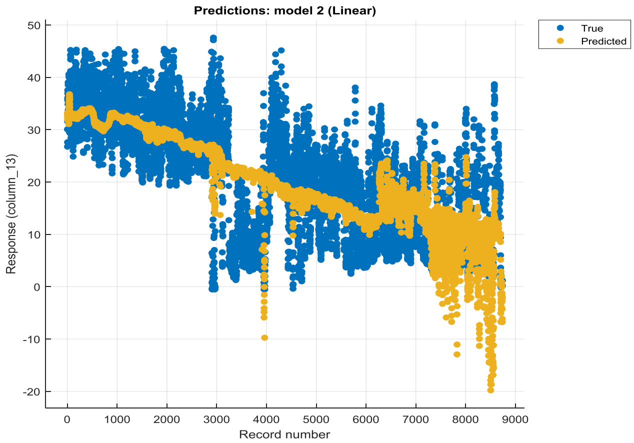


Figure 9: true and predicted data.

3 - Wide Neural Network (Model 3)

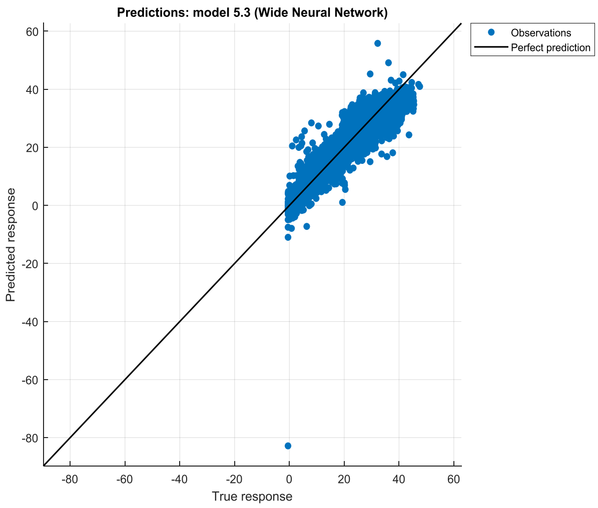


Figure 10: prediction of data.

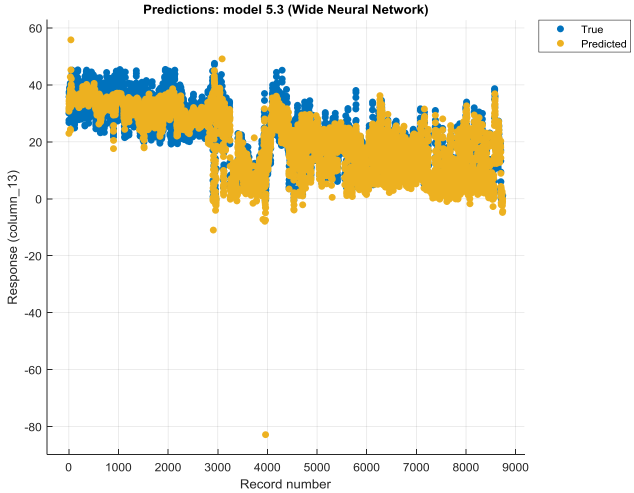


Figure 11: true and predicted data.

TABEL 3. DESCRIBE THE MEASUREMENTS OF OPTIMIZABLE MSE, R SQUARE, RMSE AND MAE

	OPTIMIZABLE DECSION TREE	OPTIMIZABLE SVM
MSE	10.86	49.6
R SQUARE	0.92	0.62
RMSE	3.29	7.04
MAE	2.225	4.121

4 - Decsion Tree (Optimize DT)

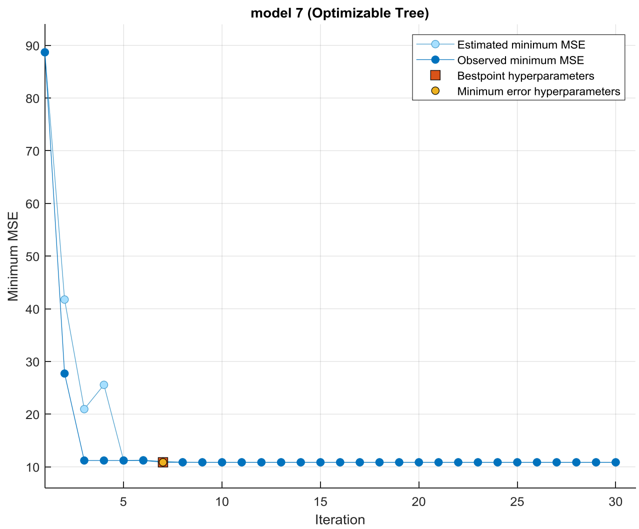


Figure 12: minimum MSE.

5 - Support Vector Machines (Optimizable SVM)

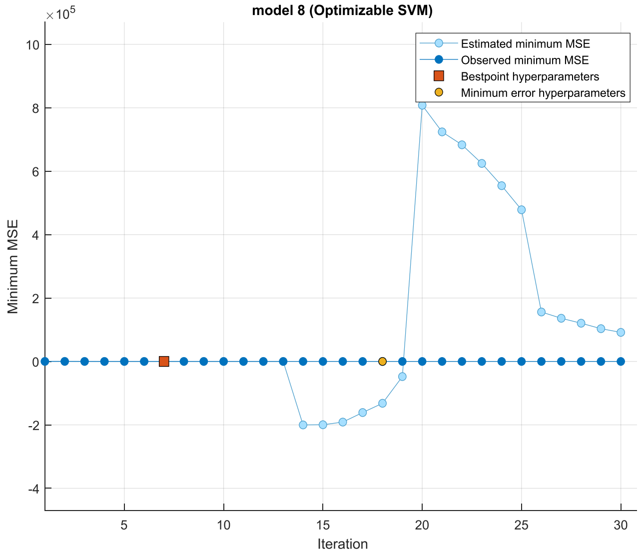


Figure 13: minimum MSE

In these models, the best result validation that is Decision Tree then we make to test of this model and the result show in the follow figure

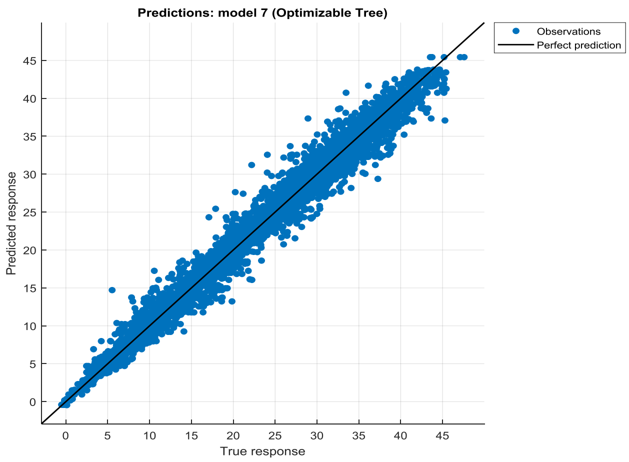


Figure 14: prediction of data.

Test Results for testing model of Decision Tree :

RMSE(Test) =1.244

R-Squared(Test) = 0.99

MSE(Test) =1.547

MAE(Test) =0.818

6 - Neural Network Fitting

To access to training data by a neural network we assume training data =70% and validation data= 15% and Test data=15% ،the performance plot as show in figur15 e.

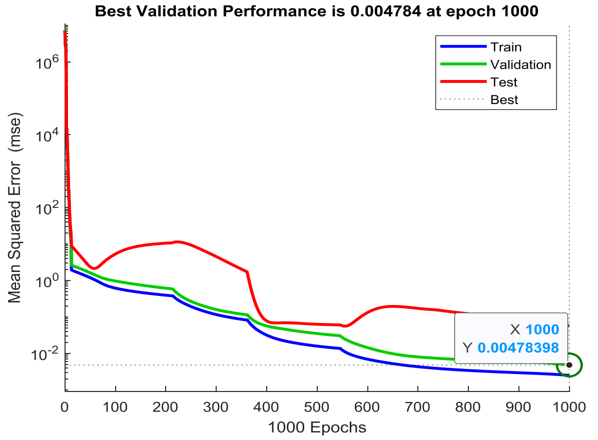


Figure 15: best validation performance.

B. *Classification Learner*

We used the same models that used in Regreation Learne and calculate Accuracy that was the best result of accuracey equal to 99.8% when applyingthe models of Decision Tree and Ensemble classifiers(Bagged Trees), plot confuion Matrix as shown in figures

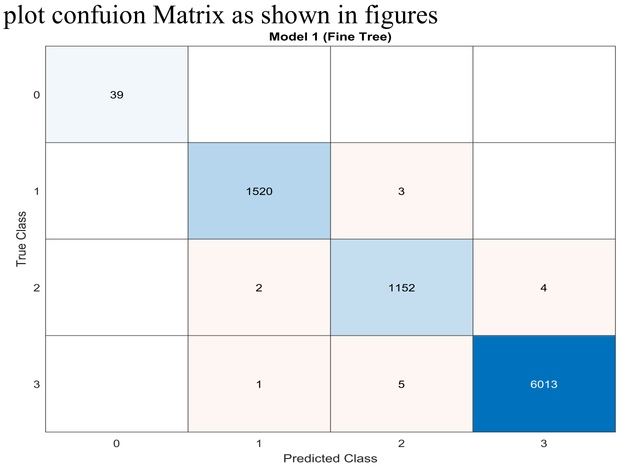
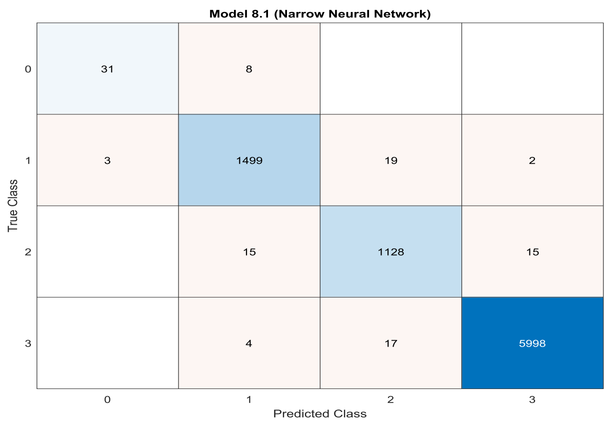
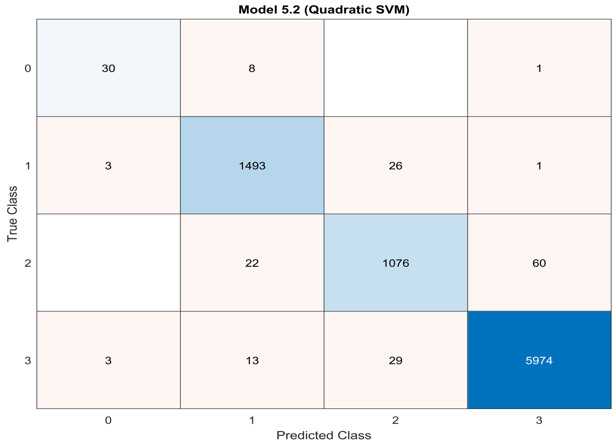


Figure 1**6**:%99.8=and Accurcey Decision Tree confuion matrix of



F7igure 1:confuion matrix of Neural Network and Accurcey=99.1%



Fconfuion matrix of SVM and Accurcey=98.1% :18 igure

we calculate of Anomaly Detection using euclidean distance for scatter plot to visulize the cluster, mean and 19 new sample as showen in follow figuer

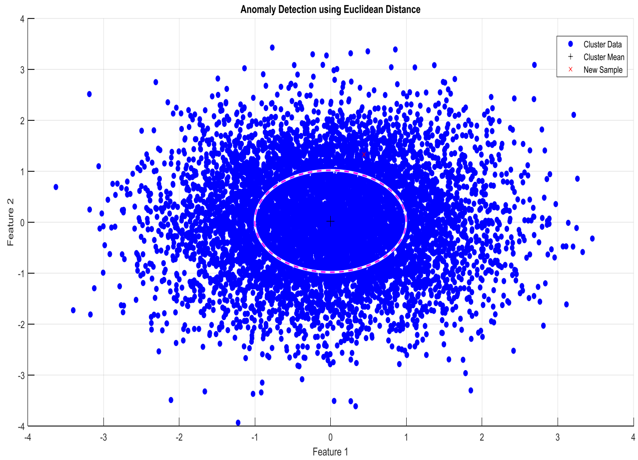


Figure 19: Anomaly detection using Euclidean distance.

CONCLUSION

Finally, in an effort to get beyond the drawbacks of conventional measuring techniques including expense, inaccuracy, and time consumption, this study investigated the use of machine learning models for well logging prediction based on well-log data. The study concentrated on assessing the effectiveness of several machine learning models, including decision trees (DT), support vector machines (SVM), Neural Network (NN) and linear regression, using characteristics like gamma ray, bulk density, neutron porosity and resistivity. The decision trees (DT) model beat the other models during the assessment process, as evidenced by its mean squared error (MSE) value of 10.86, root mean square error (RMSE) value of 3.29, mean absolute error (MAE) value of 2.225, and R-squared value of 0.92.

The best result of accuracyin classification learner equal to 99.8% when applyingthe models of Decision Tree ،(and Ensemble classifiers(Bagged Trees**.**

Based on the available well-log data, these findings demonstrate how effectively the DT model predicts well logging.

The study's conclusions demonstrate the potential of machine learning models as effective instruments for forecasting well logging and determining which well-log data makes for the best training. The industry may gain from increased well logging forecast accuracy, efficiency, and cost-effectiveness by utilizing these models, which will eventually enhance decision-making procedures and maximize resource.

REFERENCES

[1] Joshi, Dev, et al. "Prediction of sonic log and correlation of lithology by comparing geophysical well log data using machine learning principles." *GeoJournal*, vol. 1, pp. 1-22, August 2021.

[2] Talebkeikhah, Mohsen, Zahra Sadeghtabaghi, and Mehdi Shabani. "A comparison of machine learning approaches for prediction of permeability using well log data in the hydrocarbon reservoirs." *Journal of Human, Earth, and Future*, vol. 2, pp.82-99, February 2021.

[3] Ahmadi, Mohammad Ali, and Zhangxing Chen. "Comparison of machine learning methods for estimating permeability and porosity of oil reservoirs via Petro-physical logs." *Petroleum*, vol. 5, pp. 271-284, September 2019.

[4] Nordloh, Vito Alexander, Anna Roubícková, and Nick Brown. "Machine learning for gas and oil exploration." *arXiv preprint arXi*, vol. 1, pp. 3009-3016, 2020.

[5] Rahmati, Mehdi, Ghasem Zargar, and Abbas Ayati Zadeh Tanha. "Predicting density log from well log using machine learning techniques and heuristic optimization algorithm: A comparative study." *Petroleum Research* (2024).

[6] Ullah, Jar, et al. "A multidisciplinary approach to facies evaluation at regional level using well log analysis, machine learning, and statistical methods." *Geomechanics and Geophysics for Geo-Energy and Geo-Resources*, vol. 9, pp. 152, 2023.

[7] Pan, Ke-jia, Yong-ji Tan, and Hong-ling Hu. "Mathematical model and numerical method for spontaneous potential log in heterogeneous formations." *Applied Mathematics and Mechanics*, vol. 30, pp. 209-219, 2009.

[8] Zhang, Dongxiao, C. H. E. N. Yuntian, and M. E. N. G. Jin. "Synthetic well logs generation via Recurrent Neural Networks." *Petroleum Exploration and Development*, vol. 45, pp. 629-639, 2018.

[9] Struyk, C., et al. "LAS Version 2.0: A Digital Standard for Logs, Update February 2017." *Update* ,2017.

[10] Khisamutdinov, A. I., and B. V. Banzarov. "Mathematical modeling problems connected with pulsed neutron-gamma logging." *Mathematical Models and Computer Simulations*, vol. 7, pp. 79-93, 2015.

[11] Kerimova, K. A. "Classification of oil-gas bearing targets based on their saturation features and facies composition applying mathematical techniques and integrated well logging data." *Нефтегазовое дело*, vol.17, pp. 48-55, 2019.

[12] Gąsior, Irena, and Anna Przelaskowska. "Estimating thermal conductivity from core and well log data." *Acta Geophysica*, vol.62, pp. 785-801, August 2014.

[13] Rashidi, Hooman H., et al. "Common statistical concepts in the supervised Machine Learning arena." *Frontiers in Oncology*, vol.13, pp. 113-229, February 2023.