

Applied Machine Learning Methods for Detecting Fractured Zones by Using Petrophysical Logs

Haleh Azizi*, Hassan Reza

School of Electrical Engineering and Computer Sciences, University of North Dakota, Grand Forks, ND, USA
Email: *haleh.azizi@UND.edu, hassan.Reza@UND.edu

How to cite this paper: Azizi, H. and Reza, H. (2021) Applied Machine Learning Methods for Detecting Fractured Zones by Using Petrophysical Logs. *Intelligent Control and Automation*, 12, 44-64.
<https://doi.org/10.4236/ica.2021.122003>

Received: March 13, 2021

Accepted: May 14, 2021

Published: May 17, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In the last decade, a few valuable types of research have been conducted to discriminate fractured zones from non-fractured ones. In this paper, petrophysical and image logs of eight wells were utilized to detect fractured zones. Decision tree, random forest, support vector machine, and deep learning were four classifiers applied over petrophysical logs and image logs for both training and testing. The output of classifiers was fused by ordered weighted averaging data fusion to achieve more reliable, accurate, and general results. Accuracy of close to 99% has been achieved. This study reports a significant improvement compared to the existing work that has an accuracy of close to 80%.

Keywords

Decision Tree, Deep Learning, Ordered Weighted Averaging, Random Forest, Support Vector Machine

1. Introduction

Fractures play a significant role in fluid flow and oil production. Therefore, in the fractured reservoirs, e.g. Asmari Formation that is a thick sequence of shallow-water carbonates of the Zagros Basin, fractured zones should be identified and modelled. ILs, with a resolution of about 0.1 mm, are well-known tools for fracture identification, but they are usually not available. Seismic sections, well test, mud lost, petrophysical logs (PLs), and core description which have been used for fracture identification [1] [2], were associated with serious shortcomings. Seismic sections have low resolution (10 meters or less) and could be used just for the detection of faults or thick fractured zones (FZs). The resolution of

well tests and mud loss is also low. In addition, well tests are rarely available. Cores are usually unavailable, or not oriented, therefore their interpretation suffers from various shortcomings. PLs are usually available, and their resolution is about 15 cm. Recent research was focused on the identification of fractures using PLs. Researchers have selected the wells and both PLs and ILs were run inside. FZs are known by interpretation of ILs. They have applied data mining/machine learning techniques to find the relationship between fractured/non-fractured zones (F/NFZs) and PLs. In this abstract, a two-class problem must be solved.

2. Background on Machine Learning Approaches

Nobody can answer the question about which machine learning algorithm is better for a specific problem. It can only be found by applying different algorithms individually and comparing the results to figure out which one could work more accurately.

Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), and Deep Learning (DL), are four classifiers that were used in the current paper. The flowchart of the procedure was followed up for FZ detection and it is shown in **Figure 1**. Utilized methods are introduced briefly in the following sections.

2.1. Data Preprocessing

Data preprocessing is the first step in data mining. For this purpose, some data corrections such as missed data, and tools error have been conducted for all PLs.

For two classifiers SVM and DL, data needs to be normalized. Relations 1 and 2 are the ways for normalization for SVM and DL respectively [3]:

$$\hat{x}_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}, \quad \hat{x}_i \rightarrow [0,1] \quad (1)$$

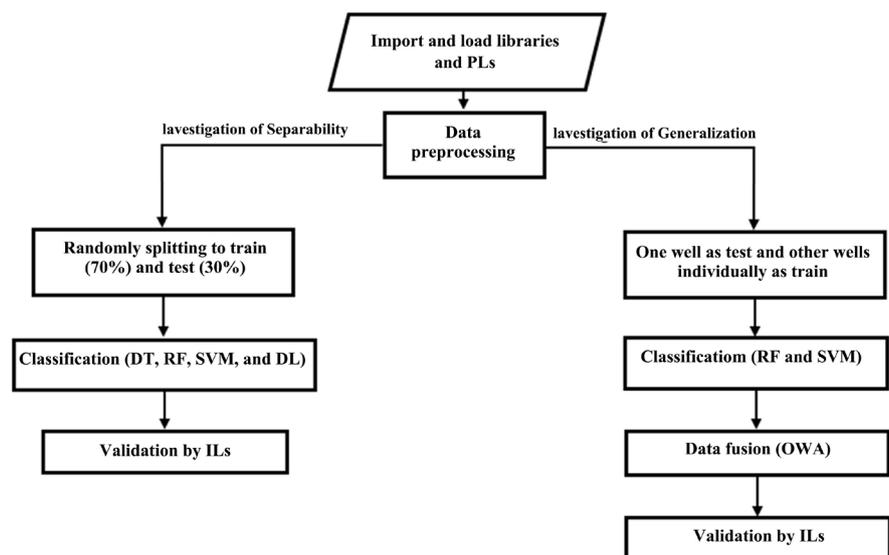


Figure 1. Steps for finding fractured zones by ML approaches.

where \hat{x}_i is the normalized value of x_i which is a well-log value, *i.e.* the registered value by an in-well sensor at a specific depth. x_{\min} and x_{\max} are minimum and maximum of well-log value, respectively.

$$\hat{x}_i = 2\hat{x}_i - 1, \hat{x}_i \rightarrow [-1, 1] \quad (2)$$

where \hat{x}_i, \hat{x}_i are the normalized values, with two different ranges.

2.2. Classifiers

The following four classifiers, used in this paper are introduced, here.

2.2.1. Decision Tree

The construction of DT is like a tree that has a root node and several interior nodes named internal nodes or decision nodes. Internal nodes include a series of questions pertaining to the data to reach a leaf node or terminal nodes which are decision nodes [4] [5].

There are several steps in order to reach a final decision in leaf nodes. First, split the data into multiple subsets according to the rule associated with the variable, in the root node. Next, split the data again in internal nodes or decision nodes recursively according to the best decision to identify the variables and the rule associated with the variable. Then, repeat the last step on the sub-nodes until reaching a stopping condition. Finally, leaf nodes are the final decisions based on the majority class label for classification goals or consider the average of the target variables which presents a regression task.

There are many algorithms for DT, such as ID3, C4.5, CART, and SLIQ. All of them work the same but the way of splitting is different.

The best splitting node in DT classification refers to obtaining homogenous as possible sub-nodes or child nodes upon splitting a parent node. This means the lower variance in sub-nodes is the better splitting in the case of regression. Classification Error, Gini Index, and Entropy are exemplar approaches that were proposed to achieve this goal. Since these approaches show the impurity of a node, a lower value of means higher homogeneity of the node.

In this research, the Gini Index approach is used. Gini is an index of the number of random data points being misclassified. This index varies between 0 and 0.5. The lower Gini Index shows the lesser chance of any random data point that is misclassified, and it helps with better decision making with lower ambiguity [4] [5].

$$G = \sum_{i=1}^k p_i (1 - p_i) = 1 - \sum_{i=1}^k p_i^2 \quad (3)$$

in which p_i corresponds to the probability of the data point belonging to i th class label and k accounts for different class labels.

2.2.2. Random Forest

RF is combining several DTs that were trained by different sets of observations. The final prediction is made by averaging the prediction of each tree. This is the

benefit of the RF rather than the DT because it solves overfitting of the training data. This also leads to higher accuracy in comparison with the DT. RF could also be useful for ranking the features. Steps of the RF are as follows. First, select the samples from the dataset. Next, create a DT and predict the results for each selected sample. Then, rank all predicted results individually by using the dataset mode factor. Lastly, find the final result by using majority vote over predicted results [4] [5].

2.2.3. Deep Learning

A multilayer-perception (MLP) is a deep, artificial neural network; it is an entrance gate to the DL world. The structure of MLP consists of an input layer to receive a raw dataset and the last layer which is an output layer that makes the final decision. Between these two layers (input/output layers), there is an arbitrary number of hidden layers that are the core of the MLP and consist of functions and weighting coefficients that belong to them [4] [5].

Multi-layer perceptions train on pairs of input and output layers to discover the best model. Simultaneously, in training model functions, weighting coefficients, and biases are optimized to minimize the error between result data in the output layer and the test data. Backpropagation minimizes the error. The error also can be measured in a variety of ways, for instance, root mean squared error (RMSE).

In this paper the number of features was 10, the number of hidden layers was 3 in different nodes. Also, the activation function was ReLu, and the model was iris. DL was run up to epoch or step 200. The first assumption is a matrix, as an input dataset by n sample where each has features. The second assumption is that MLP has hidden layers. The third assumption is that the hidden layers have weights and biases and the fourth assumption is that the output layer has biases. Therefore, the outputs of the one hidden layer MLP is as follows [4] [5]:

$$\begin{aligned} H &= XW^{(1)} + b^{(1)}, \\ O &= HW^{(2)} + b^{(2)}. \end{aligned} \quad (4)$$

A nonlinear activation function σ is applied to each hidden node in which outputs of activation functions are called activations. So, multi-layer architecture will be as follows [4] [5]:

$$\begin{aligned} H &= \sigma(XW^{(1)} + b^{(1)}), \\ O &= HW^{(2)} + b^{(2)}. \end{aligned} \quad (5)$$

By stacking hidden layers, the next hidden layer will define one another as follows [4] [5]:

$$\begin{aligned} H^{(1)} &= \sigma_1(XW^{(1)} + b^{(1)}), \\ H^{(2)} &= \sigma_2(H^{(1)}W^{(2)} + b^{(2)}). \end{aligned} \quad (6)$$

The most popular activated functions are the rectified linear unit (ReLU), sigmoid function or squashing function, and the tanh (hyperbolic tangent) func-

tion.

2.2.4. Support Vector Machine

SVM is one of the most powerful and flexible methods in machine learning and can be used for both classification and regression; however, it is recommended to be utilized as a classifier. The important concepts in SVM are support vectors, hyperplane, and margin.

To find a better classifier by SVM, researchers should find the maximum marginal hyperplane (MMH). Therefore, SVM will generate several hyperplanes iteratively for the best discrimination and this method chooses the best hyperplane that works correctly [4] [5]. Kernels in the SVM classifier convert a non-dimensional dataset to a separable space by adding more dimensions to it. Briefly, the SVM classifier is implemented with a kernel that transforms input data into the higher dimensional space with the best class discrimination.

Linear, Polynomial, and Radial Basis Function (RBF) are popular kernels that are used in SVM. F/NFZs are not linearly separable; therefore, RBF is the kernel that was used in the current paper. The RBF kernel is a kind of exponential kernels that is common in SVM classification and maps input space in definite dimensional space. The following formula explains it mathematically [4] [5]:

$$K(x, x_i) = \exp\left(-\text{gamma} * \sum (x - x_i)^2\right) \quad (7)$$

in which gamma is in the range of 0 to 1.

SVM classifier is an effective method when the number of features is quite large. It is a powerful classifier model since it maximizes margin.

After applying classifiers, the question is which result is the best one? Ambiguity in decision-making leads to data fusion. Integrating the results achieved from different classifiers and all wells, to achieve more reliable, robust, and accurate discrimination is the goal of Ordered Weighted Averaging Data Fusion.

2.3. Data Fusion

Ordered weighted averaging (OWA) was introduced by Ronald R. Yager (1988) [6]. This operator is commonly used in the decision-making processes. It is still a powerful and yet simple method for finding aggregate output from data series. The OWA operator with n dimension is a mapping $F: R^n \rightarrow R$ and has an associated n vector. Here n is the number of inputs that must be fused to achieve a final result. In the current study, n is equal to 14, because the database of seven wells times two classifiers (RF and SVM) were applied to discriminate FZs from NFZs. Weights are in the range of zero and one ($w_i \in [0,1]$), and their summation have to be equal to 1 [6]:

$$\sum_{i=1}^n w_i = 1 \quad (8)$$

In each depth, results achieved from different training wells and classifiers are defined as a_k . So, a_k could be zero (label which represents NFZs) or one (label which represents FZs). The operator works as follows:

$$F(a_{k_1}, a_{k_2}, \dots, a_{k_n}) = \sum_{j=1}^n b_{k_j} w_j = b_{k_1} w_1 + b_{k_2} w_2 + \dots + b_{k_n} w_{j_n}, k \quad (9)$$

b_{k_j} are in fact a_k which have a descending order. Therefore, b_1 is the biggest label (usually 1) and b_n the least (usually 0). k represents the depths of the test well. Here, output of this cumulative operator is always in the range of zero and one, and could be said in the group of "OR" and "AND" [7] [8]:

$$\min_i [a_i] \leq F_w(a_1, a_2, \dots, a_n) \leq \max_i [a_i] \quad (10)$$

Yager introduced an Orness evaluation that is related to the vector addressed by w [7]:

$$\text{orness}(w) = \frac{1}{n-1} \sum_{i=1}^n (n-i) w_i \quad (11)$$

Therefore, for every " w ", the orness (w) is located at a unit distance. Suppose that d_k represents the real fracture label in each depth of a test well. The error should be minimized by optimizing the weights (w_i) [7]:

$$e = 1/2 \left[\sum_{k=1}^m (b_{k_1} w_1 + b_{k_2} w_2 + \dots + b_{k_n} w_n - d_k)^2 \right] \quad (12)$$

The main problem is that the OWA operator must act everywhere in a way that the error is minimized. In other words, weights (w_i) must count in such a way that in the end, the error (e) goes towards zero. This number should be obtained while the solution is optimized. Yager has proposed two scenarios for optimization of the weights: optimistic and pessimistic.

In optimistic ordered weighted averaging (OOWA), weights (w_i) are defined as a function of coefficient (α) [6]:

$$\begin{aligned} w_1 &= \alpha; w_2 = \alpha(1-\alpha); w_3 = \alpha(1-\alpha)^2; \dots; \\ w_{n-1} &= \alpha(1-\alpha)^{n-2}; w_n = (1-\alpha)^{n-1}; 0 \leq \alpha \leq 1 \end{aligned} \quad (13)$$

Therefore, it is enough that α varies from zero to one, and error (e) associated to a different α is calculated. Minimum error corresponds to optimum α . It could be said that optimum weights (w_i) in OOWA.

In pessimistic ordered weighted averaging mechanism (POWA), weights (w_i) again are defined as a function of the α [6]:

$$\begin{aligned} w_1 &= \alpha^{n-1}; w_2 = (1-\alpha)\alpha^{n-2}; w_3 = (1-\alpha)\alpha^{n-3}; \dots; \\ w_{n-1} &= (1-\alpha)\alpha; w_n = (1-\alpha); 0 \leq \alpha \leq 1 \end{aligned} \quad (14)$$

The procedure for optimization of weights (you say α) is similar to the optimistic mechanism.

3. Related Works

Mazaheri and her associates used Artificial Neural Network to estimate Fracture Measure (FM), as a parameter which was calculated by aperture, fracture type, azimuth, and apparent distance [9]. Their results that pertained to finding the relation between FM and conventional logs were satisfactory with a generalized

correct classification rate (CCR) of about 80% [10]. They optimized cell size to more effectively utilize FM to recognize FZs. They applied different data mining techniques to do that and fused the results with the Sugeno integral method. Comparison between the results confirmed that 30 cm is the optimum cell size [11]. Tran studied characterization and modeling of naturally fractured reservoirs. Tran also tried to identify the fractures as a data source [12]. He integrated and classified different logs to discriminate FZs from NFZs. He could not validate the results because he did not access ILs.

Olivia *et al.* developed a fast region-based Convolutional Neural Network (fast-RCNN) for automatic interpretation of acoustic image logs (ILs) to identify fractures and breakouts. They achieved around 81% accuracy in fracture detection which is satisfactory [13]. However, it should be mentioned that they just automatically interpreted ILs instead of manually interpreting them. Therefore, their aim was not to develop a method to discriminate FZs by using PLs.

Tokhmechi and his associates, decomposed water saturation by wavelet transformation and showed that the majority of the information from the original log is hidden at low-frequency bands. The approximate section of water saturation of filtered zones helped them to detect FZs with an accuracy of about 70% [14]. They, in another research study, applied Parzen and Bayesian classifiers to raw, de-noised, and various frequency bands of PLs for FZ detection [15]. They utilized OWA (ordered weighted averaging) to fuse the results obtained from different training wells and classifiers as well. The generalization of their approach was confirmed with an average accuracy of about 72%. Their third study was focused on fracture density estimation by utilizing wavelet and Fourier transformation to calculate the energy of the PLs in the FZs [16]. Linear and non-linear regressions were fitted between the energy of logs and fracture densities. Results showed a strong correlation between the energy of caliper, sonic, density (RHOB), and lithology (PEF) logs with fracture density in each well. Martinez-Torres utilized fuzzy logic to integrate caliper, gamma-ray, sonic, self-potential, and resistivity logs to make a composite fracture log [17]. Lack of IL for verification of the proposed approach is his research's shortcoming.

Mazhari and associates proposed a generalized Case-Based Reasoning (CBR) method for FZ detection via PLs [10]. To such an aim, they used a set of training wells to create a database composed of both petrophysical data and the ILs. They conducted a learning automata-based algorithm to find the optimal similarity relation between PLs and manual interpretation of the borehole ILs. Their developed model was successfully tested on the Asmari reservoir through several oil wells and it achieved a general CCR of about 70%.

Jafari *et al.*, studied an adaptive neuro-fuzzy inference system for fracture density estimation from PLs [18]. They found a good statistical correlation between fracture density and well log data including sonic, deep resistivity, neutron porosity, and bulk density, and achieved a correlation coefficient of 98% between the measured and neuro-fuzzy estimated fracture density.

Zarehparvar Ghoochaninejad and his colleagues estimated hydraulic aperture of detected fractures using well log responses using a Teaching–Learning–Based Optimization algorithm (TLBO), which trained an initial Sugeno fuzzy inference system [19]. Aghli *et al.*, tried to find a quick generalized method for identification of FZs using PLs and then used the Velocity Deviation Log (VDL) to identify fracture aperture opening and their effects on porosity and permeability in high fracture density zones [20]. Discrimination of lithofacies to the optimization of FZ detection using PLs was another approach to maximize the accuracy [21]. The aforementioned chained activities show the importance of using data mining techniques to discriminate FZs from NFZs, which is also the approach of the current research. Flavio and Gregor proposed a composite well log, called velocity deviation log, to estimate different reservoir properties, and fracture identification as well [22]. Yan and associates proposed a fracture-cased porosity modeling procedure [23]. They created synthetic IL by integrating PLs. They have found that the procedure is useful not only for fracture detection but also for fracture properties estimation. It seems this procedure is straight-forward and might be developed in the future.

In this paper, different classifiers were developed in Python Jupiter notebook and Google Colab, to better discriminate FZs from NFZs. The proposed classifiers containing DT, RF, SVM, and DL, were applied over eight oil wells of a giant fractured reservoir. Their PLs and ILs are available. Outputs are compared and then fused by OWA to achieve more accurate, reliable, and general results.

4. Database

The database used in this study is one of the world's giant carbonate fractured reservoirs. Exploration over the studied field was started in 1923 and the first oil well was drilled in 1931. The formation of the reservoir is Asmari which is a fractured carbonate rock. More than 450 oil wells have been drilled so far. Full set PLs were run inside almost all of the wells, while ILs were run in just eight of them. Petrophysical and ILs of eight wells were used in the current study. They are located over semi UGC map¹ of the studied reservoir (**Figure 2**).

In **Table 1**, a list of PLs that were run inside eight studied wells are reported. Overall, 29 raw and interpreted logs were available in the studied wells. For instance, density (RHOB), sonic and photoelectric factor (PEF) are examples of raw and dolomite, porosity, and summation of gamma ray (SGR) are examples of interpreted logs (**Table 1**). 16 out of 29 logs were selected for more studies.

In **Table 2**, the availability of selected logs in eight studied wells are listed. ILs were also run in all eight studied wells. Interpreted IL is the source of class labels (F/NFZS) and is also used for validation.

As can be seen in **Table 2**, some of the logs are not available in different wells. For instance, geological rock type logs, e.g., dolomite, shale, or limestone are not available in about half of the wells. Raw PLs containing caliper, cumulative

¹Underground Counter Map.

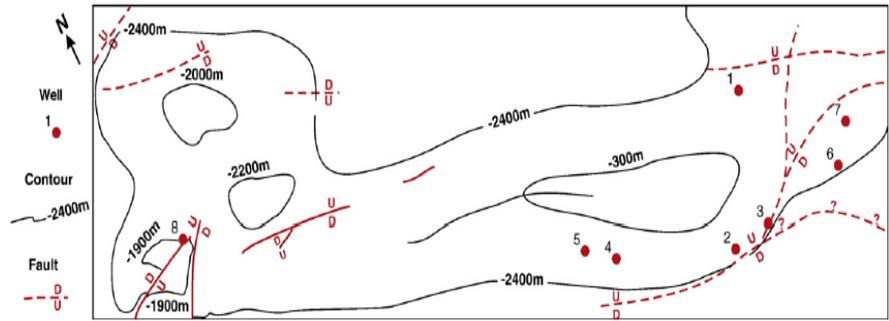


Figure 2. Semi UGC map and location of studied wells.

Table 1. Discrimination of available PLs between raw and interpreted ones.

Kind of log	Logs
Raw	Caliper, Gamma Ray (Potassium, Thorium, Uranium, GR), Resistivity (MSFL, LLS, SFL, ILM, ILD, LLD, DFL, RXO, RT), Sonic, Density (RHOB, PEF), Neutron (NPHI)
Interpreted	Lithology (Limestone, Dolomite, Sandstone, Shale, Anhydrite), Gamma Ray (SGR, CGR), Porosity and Saturation (POR, PHI, TOTAL, SEC, 2ASW × PHIO, SXO, SW)

Table 2. Availability of petrophysics logs in studied wells.

Well	1	2	3	4	5	6	7	8
Well Log								
Caliper	*	*	*	*	*	*	*	*
Geological Rock Type	Limestone	*				*	*	*
	Dolomite	*			*	*	*	*
	Sandstone				*	*	*	*
	Shale					*	*	*
	Anhydrite	*			*	*	*	*
Gamma Ray	Potassium	*	*	*		*		*
	Thorium	*	*	*		*		*
	Uranium	*	*	*		*		*
	CGR	*	*	*	*	*	*	*
	GR	*			*			
	SGR	*	*	*		*	*	*
Resistivity	MSFL				*			
	LLS (SFL)				*			*
	ILM		*				*	*
	ILD (LLD)	*	*	*	*	*	*	*
	DFL						*	
	RXO				*	*		*
	RT	*			*	*	*	*
Sonic	Sonic	*	*	*	*	*	*	*

Continued

Density	RHOB	*	*	*	*	*	*	*	*
	PEF	*	*	*	*	*	*	*	*
Neutron	NPHI	*	*	*		*	*	*	*
Porosity and Saturation	Porosity (PHI)	*			*	*	*		
	Total					*	*		
	SEC				*	*	*		
	2ASW*PHIO				*		*	*	*
	SXO				*			*	*
	SW	*			*		*	*	*

*means available.

gamma-ray (CGR), SGR, RHOB, Sonic, PEF, and neutron porosity (NPHI), are available in almost all of the wells. Water saturation (SW) and Porosity (effective “PORE” and total “PORT”) logs are also available in the majority of the wells. Resistivity logs containing MSFL, LLS, SFL, ILM, ILD, LLD, DFL, RXO, RT are not available in the majority of the wells; therefore, they will not be used in the current study.

In the next step, PLs that are helpful to FZ detection are selected. For this purpose, the correlation coefficient of PLs with FZs is used. However, this factor had a very low value (less than 0.4). Minor differences are a sign of the complexity of discrimination of FZs from non-fractured ones. Therefore, the reasons for selection were a higher correlation coefficient factor, affected logs of FZ, and availability of log in the studied well. In **Table 3**, the effects of FZs over PLs are briefly discussed.

Overall, caliper, CGR, SGR, RHOB, Sonic, PEF, NPHI, SW, effective and total porosity were selected for FZ detection. Therefore, classifiers will define discrimination F/NFZs in 10D feature space.

5. Results and Discussion

Results and discussion about them are presented as follows:

5.1. Possibility of Discrimination of Fractured from Non-Fractured Zones

Confusion matrix, accuracy or correct classification rate (CCR), have been calculated, as classifier’s performance index. Those evaluation parameters are briefly introduced as follows:

Confusion matrix (C) is a square matrix, in which entry diagonals represent the number of accurate classified data, and the rest represents misclassified data. The current study is two-class problem in which C00 represents the count of true negative, which means true classified NFZs (**Figure 3**). C01 represents false negative or NFZs which are misclassified as fractured ones. C10 represents false

Table 3. The effect of FZs on PLs.

Log	A common observation in the majority of the studied wells.
Density (RHOB)	Usually, because of the low density in the FZs toward rocks, RHOB is lower in the fractured area.
Sonic	The sonic log is based on wave arrival time. This time depends on P wave velocity that passes through lithology, porosity, and fluids. If the wave path passed a fracture zone, the time increases because P wave velocity is slower in porous and fluids rather than solids.
Gamma Ray (GR)	GR is natural radioactivity that comes from Uranium, Thorium, and Potassium that have accumulated in fractures.
Potassium	Potassium is found in Feldspar and clay minerals. Because of the lake of Feldspar in the study formation, Potassium is due to clay formation. Clay has plasticity, therefore, it indicates the NFZ.
Thorium	Claystone (shale) contains Thorium. Shale has plasticity properties and it is contrary to the properties of fractured rock.
Water Saturation (SW)	Higher water saturation usually indicates open fracture zones.
Uranium	Uranium can dissolve in water and deposit in fractures.
Resistivity (RT)	By trapping brine in fractures, resistivity decreases. It means the NFZ has a high RT.
Photoelectric Factor (PEF)	FZs in some of the study wells indicate high PEF because of the tarp of water. It is because elements of connate water have a higher atomic number than rocks.
Caliper	The Caliper tool reads the diameters of borehole more precisely when in FZ rock edge chipped away while drilling. Also, sometimes because of mud accumulation in open fractures a caliper log is even less than bit size.
Neutron Porosity (NPHI)	NPHI would increase if a fracture zone filled with fluid or Hydrogen content.
Geological Rock Types (dolomite, limestone, shale, and anhydrite)	Dolomite and limestone are brittle, with elastic behavior, capable of fracturing. Shale and anhydrite are ductile, capable of plastic behavior.

$$\begin{bmatrix} C_{00} & C_{01} \\ C_{10} & C_{11} \end{bmatrix} = \begin{bmatrix} TN & FN \\ FP & TP \end{bmatrix}$$

Figure 3. Confusion matrix, which contains true negative (TN), false negative (FN), false positive (FP), and true positive (TP).

positive or FZs which are misclassified as NFZs. Finally, C11 represents true positive, or the FZs are correctly classified [24].

Accuracy or Correct Classification Rate (CCR) represents the number of correctly classified data divided by total number of the data [24]:

$$CCR = \frac{TN + TP}{TN + FP + TP + FN} \tag{15}$$

In **Table 4**, DT, RF, DL, and SVM classifiers are completed for all PL date in the eight wells. The Confusion Matrix and CCR are the results of these classifications. The accuracy that represents the number of correctly classified data instances

Table 4. The results of discrimination of FZs from NFZs using DT, RF, DL, and SVM.

Well	DT	RF	DL	SVM
1	$\begin{bmatrix} 318 & 44 \\ 52 & 234 \end{bmatrix}$, CCR = 0.85	$\begin{bmatrix} 342 & 20 \\ 39 & 247 \end{bmatrix}$, CCR = 0.91	$\begin{bmatrix} 315 & 29 \\ 30 & 274 \end{bmatrix}$, CCR = 0.91	$\begin{bmatrix} 311 & 33 \\ 24 & 280 \end{bmatrix}$, CCR = 0.91
2	$\begin{bmatrix} 454 & 33 \\ 53 & 378 \end{bmatrix}$, CCR = 0.91	$\begin{bmatrix} 466 & 28 \\ 30 & 394 \end{bmatrix}$, CCR = 0.94	$\begin{bmatrix} 402 & 75 \\ 61 & 380 \end{bmatrix}$, CCR = 0.85	$\begin{bmatrix} 452 & 24 \\ 24 & 418 \end{bmatrix}$, CCR = 0.95
3	$\begin{bmatrix} 328 & 19 \\ 9 & 79 \end{bmatrix}$, CCR = 0.94	$\begin{bmatrix} 338 & 9 \\ 15 & 73 \end{bmatrix}$, CCR = 0.94	$\begin{bmatrix} 318 & 29 \\ 19 & 69 \end{bmatrix}$, CCR = 0.89	$\begin{bmatrix} 325 & 17 \\ 10 & 83 \end{bmatrix}$, CCR = 0.94
4	$\begin{bmatrix} 532 & 10 \\ 9 & 62 \end{bmatrix}$, CCR = 0.97	$\begin{bmatrix} 539 & 3 \\ 12 & 59 \end{bmatrix}$, CCR = 0.98	$\begin{bmatrix} 520 & 14 \\ 19 & 61 \end{bmatrix}$, CCR = 0.95	$\begin{bmatrix} 518 & 16 \\ 17 & 62 \end{bmatrix}$, CCR = 0.95
5	$\begin{bmatrix} 308 & 10 \\ 21 & 185 \end{bmatrix}$, CCR = 0.94	$\begin{bmatrix} 309 & 9 \\ 15 & 191 \end{bmatrix}$, CCR = 0.95	$\begin{bmatrix} 296 & 21 \\ 28 & 175 \end{bmatrix}$, CCR = 0.90	$\begin{bmatrix} 303 & 9 \\ 20 & 192 \end{bmatrix}$, CCR = 0.95
6	$\begin{bmatrix} 427 & 32 \\ 22 & 373 \end{bmatrix}$, CCR = 0.94	$\begin{bmatrix} 442 & 17 \\ 20 & 375 \end{bmatrix}$, CCR = 0.96	$\begin{bmatrix} 357 & 89 \\ 90 & 311 \end{bmatrix}$, CCR = 0.79	$\begin{bmatrix} 437 & 13 \\ 16 & 388 \end{bmatrix}$, CCR = 0.97
7	$\begin{bmatrix} 286 & 9 \\ 10 & 172 \end{bmatrix}$, CCR = 0.96	$\begin{bmatrix} 279 & 16 \\ 9 & 173 \end{bmatrix}$, CCR = 0.95	$\begin{bmatrix} 264 & 32 \\ 19 & 163 \end{bmatrix}$, CCR = 0.89	$\begin{bmatrix} 264 & 6 \\ 10 & 197 \end{bmatrix}$, CCR = 0.97
8	$\begin{bmatrix} 414 & 0 \\ 6 & 262 \end{bmatrix}$, CCR = 0.99	$\begin{bmatrix} 407 & 7 \\ 18 & 250 \end{bmatrix}$, CCR = 0.96	$\begin{bmatrix} 354 & 55 \\ 75 & 193 \end{bmatrix}$, CCR = 0.81	$\begin{bmatrix} 400 & 2 \\ 6 & 274 \end{bmatrix}$, CCR = 0.99

over the total number of data instances shows in this table that all classifiers work accurately more than at least 85%. To investigate which methods are the best **Figure 4** is shown.

In **Figure 4**, CCR for different classifiers in all studied wells, as well as the average CCR are presented. Comparison between classifiers shows that DL has the weakest discrimination, while both RF and SVM have the best. Based on the results, RF and SVM are the two selected classifiers for the next step, which is using the database of a well to discriminate FZs from NFZs in other wells. On the horizontal axis, the average of all classifiers in all wells is shown.

5.2. Generalization of Classification

In this section, for Generalization, one well data is considered as a test, and the data of the other 7 wells is used as training data. This is for checking the possibility of using trained classifiers for fracture detection in other wells. RF and SVM are the classifiers that were used for this goal. The results are shown in **Table 5**.

In **Table 5**, CCRs, which are the result of classification when one well is test and others were utilized for training are presented. For example, in **Table 5**, column two, wells 1, 3, 4, 5, 6, 7, and 8 were trained individually using the classifier (RF and SVM) to discriminate F/NFZs, and then trained classifier was tested in well 2. As it shows, CCR is equal to 98% for all runs.

To investigate more using the generalized classification, how much will the result improve? **Figure 5** was created to answer this question. In this figure average CCRs (ACCR) for each test well, while RF or SVM that were utilized for classification are displayed. The amazing point is that in all cases ACCRs is

Table 5. Results of discrimination between F/NFZs using RF and SVM when one well is test and another well is train*.

TW	Test Well # 1			Test Well # 2			Test Well # 3			Test Well # 4			Test Well # 5			Test Well # 6			Test Well # 7			Test Well # 8		
	RF	SVM	TW	RF	SVM																			
2	0.96	0.97	1	0.98	0.98	1	0.98	0.98	1	1	0.98	1	0.98	0.98	1	0.99	0.99	1	0.99	0.98	1	0.99	0.99	
3	0.97	0.97	3	0.98	0.98	2	0.97	0.98	2	0.99	0.99	2	0.98	0.98	2	0.99	0.99	2	0.99	0.99	2	0.99	0.99	
4	0.97	0.96	4	0.98	0.98	4	0.97	0.97	3	0.99	0.98	3	0.98	0.98	3	0.98	0.99	3	0.98	0.99	3	1	0.99	
5	0.96	0.97	5	0.99	0.98	5	0.98	0.98	5	0.99	0.99	4	0.98	0.98	4	0.99	0.99	4	0.99	0.99	4	0.99	0.99	
6	0.97	0.97	6	0.98	0.98	6	0.98	0.97	6	0.99	0.98	6	0.98	0.98	5	0.99	0.99	5	0.98	0.99	5	0.99	0.99	
7	0.96	0.97	7	0.99	0.98	7	0.98	0.98	7	0.99	0.99	7	0.98	0.98	7	0.98	0.98	6	0.98	0.98	6	0.99	0.99	
8	0.97	0.96	8	0.98	0.98	8	0.98	0.98	8	0.99	0.99	8	0.97	0.97	8	0.99	0.98	8	0.98	0.99	7	0.99	1	

*TW: Training Well.

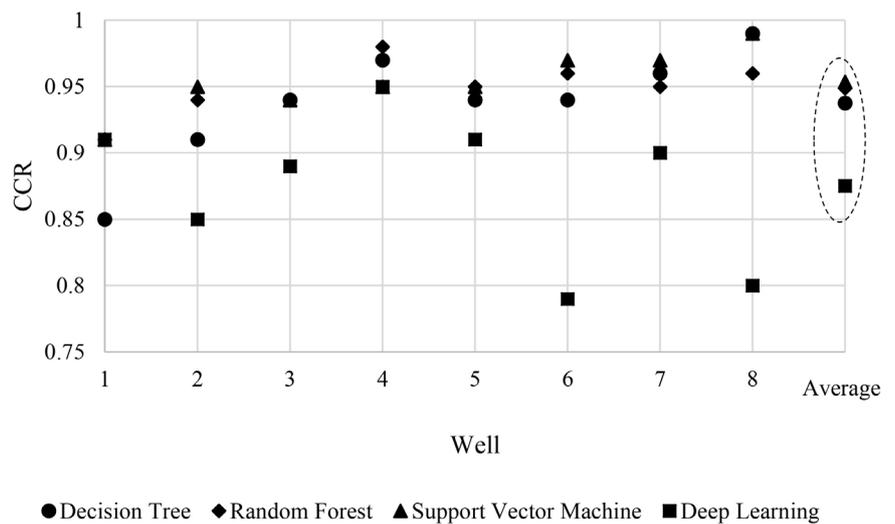


Figure 4. Comparison between average correct classification rate (CCR) of discrimination of F/NFZs by using different classifiers over studied wells.

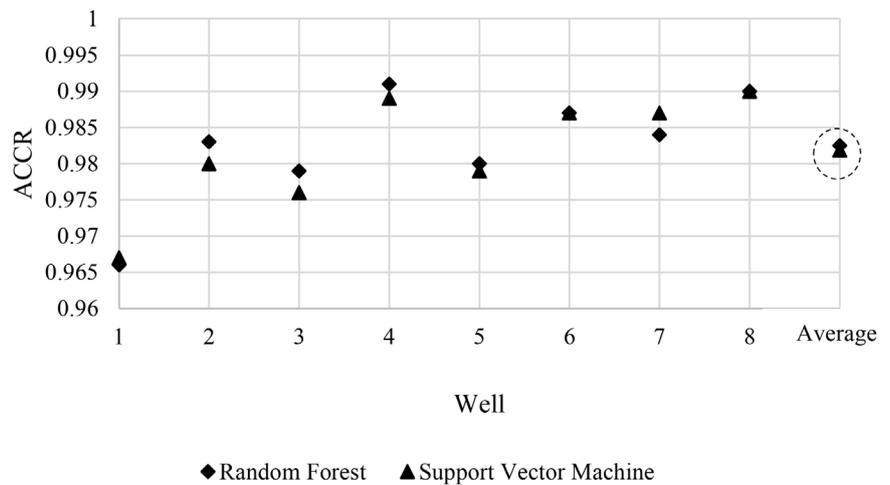


Figure 5. Comparison between average correct classification rate (ACCR) of discrimination of F/NFZs by using RF and SVM classifiers over studied wells.

higher than 96 percent, while in previous studies, ACCRs using Parzen, Bayesian [15], and case based reasoning [10] classifiers reported less than 70 percent. Therefore, RF and SVM, are great; however, the performance of RF is a little bit better than SVM (average in Figure 5).

The best tool for validation of the result of this paper is ILs that have FZ information with a resolution of millimeters. This will be obtained when the results (FZ) extract from classifiers in any well are compared to ILs that belong to the same well.

In Figure 6 and Figure 7 discrimination between F/NFZs for wells 1, while using RF and SVM classifiers are displayed, and could be comprised by real zones achieved in the interpretation of ILs.

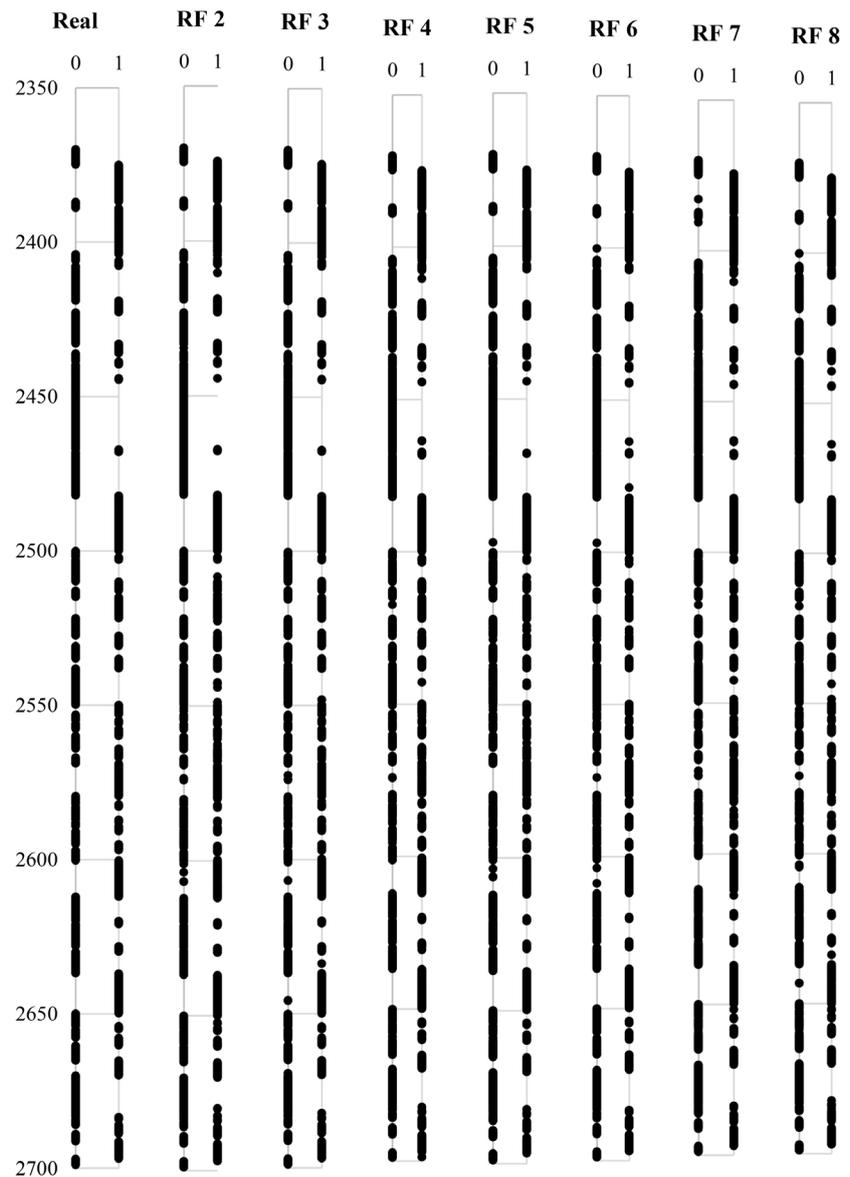


Figure 6. Discrimination between FZs and NFZs for well #1, when PLs of other wells (wells 2 through 8) were used for training of RF classifier (0: NFZ and 1: FZ).

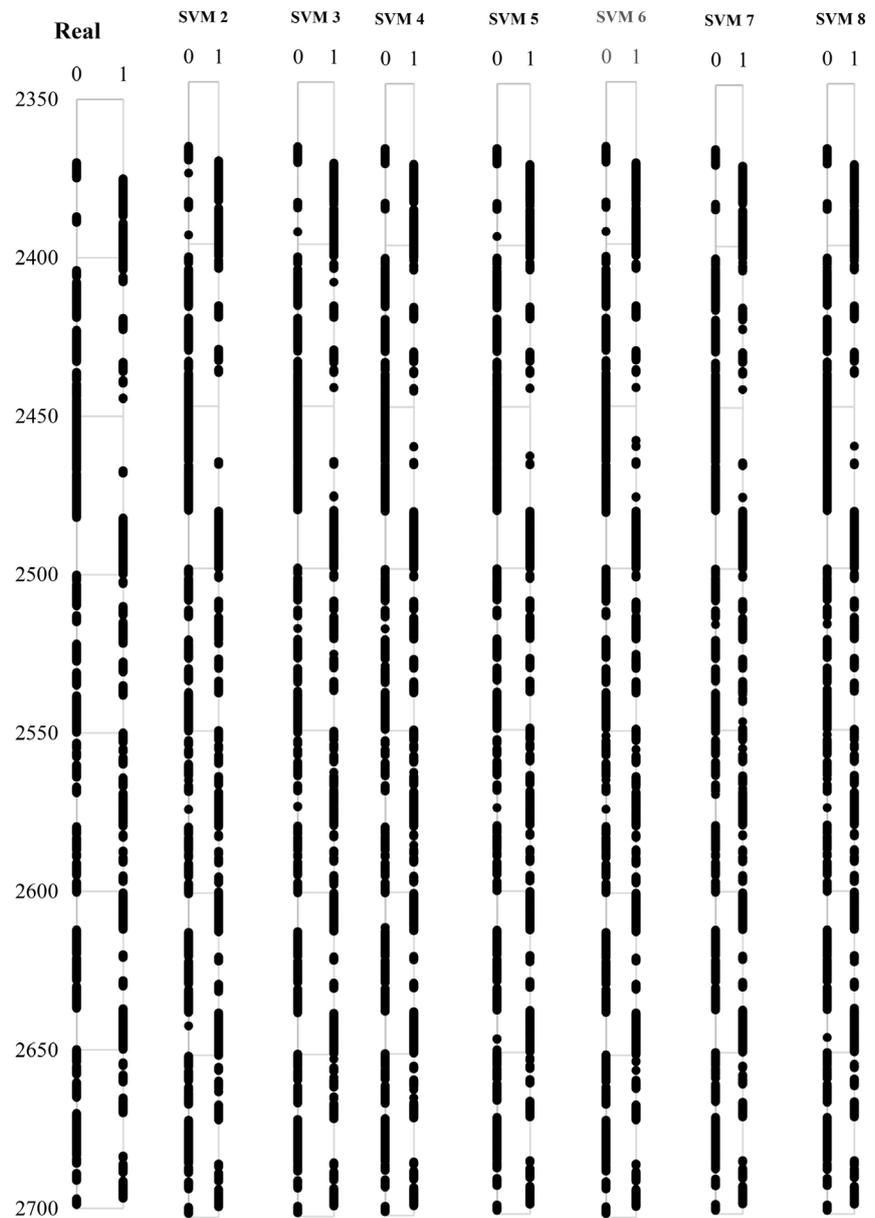


Figure 7. Discrimination between FZs and NFZs for well #1, when PLs of other wells (wells 2 through 8) were used for training of SVM classifier (0: NFZ and 1: FZ).

In **Figure 6** and **Figure 7**, FZs are labeled with 1 and zero for NFZs. These figures confirm the validation of the results. However, by digging into the results, it was observed that errors occurred in the boundaries of F/NFZs. In the industry, FZ recognition is much more important than the resolution of the boundaries; therefore, the achieved results are highly encouraging.

Now, the question is for the wells in which IL is not available. Which well should be considered for training? Ambiguity in decision-making leads to data fusion. Integrating the results achieved from different classifiers and all wells, to achieve more reliable, robust, and accurate discrimination is the goal of data fusion. To overcome this problem, OWA was selected for fusion. Optimistic

(OOWA) and pessimistic (POWA) scenarios were applied over the 14 results of classifiers, and optimum which minimizes the sum of squared error (SSE), was found. It should be emphasized that SSE in the current study is equal to the number of misclassified data. In other words, SSE is equal to false negative (FN) plus false positive (FP) in **Figure 2**.

Results of optimization (Minimum SSE) for wells 3 and 4 as examples are presented in **Figure 8**. As can be seen, accuracy has increased to about 100%. Amazing results, in comparison to the best results (about 72%) that have been published so far [10] [15].

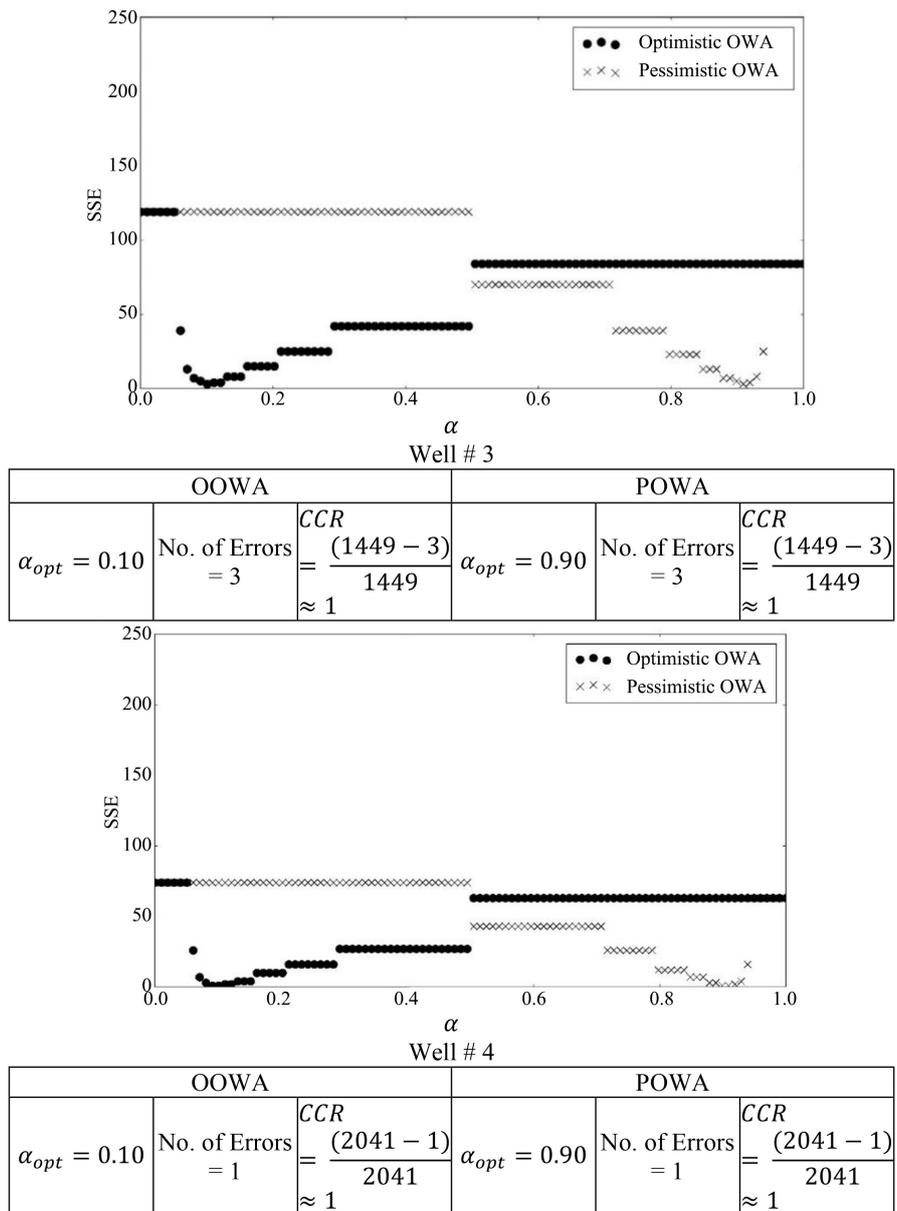


Figure 8. 14 results of classifications using training data were fused using OWA data fusion method in two scenarios (optimistic and pessimistic). The results of optimization of α , number of errors, and approximate CCR for wells 3 and 4 are presented.

It should be emphasized that in all cases SSE for OOWA and POWA is equal together (Table 6); therefore, there is no priority for one of them. α_{opt} differs in the range of 0.09 to 0.13 in optimistic and 0.89 to 0.92 in pessimistic scenarios. In general, α_{opt} could be considered equal to 0.11 and 0.91 for optimistic and pessimistic scenarios respectively.

An average number of errors (FP plus FN) for RF, SVM, and OWA Table 7 are reported. This Table helps to rank the methods and kind of training database.

Table 7 also highlights that data fusion has improved the results by more than 85 percent. Approximately, the summation of the number of errors for data fusion declined from about 300 to 40.

ROC (Receiver Operating Characteristic) curve is a fundamental graphical tool evaluation for a range of diagnostic test result.

Table 6. OWA data fusion method, in two scenarios (optimistic and pessimistic) for all study wells.

		OOWA		POWA		
Well # 1	$\alpha_{opt} = 0.12$	No. of Errors = 14	$CCR = \frac{2159-14}{2159} \approx 1$	$\alpha_{opt} = 0.91$	No. of Errors = 14	$CCR = \frac{2159-14}{2159} \approx 1$
Well # 2	$\alpha_{opt} = 0.12$	No. of Errors = 4	$CCR = \frac{3057-4}{3057} \approx 1$	$\alpha_{opt} = 0.91$	No. of Errors = 4	$CCR = \frac{3057-4}{3057} \approx 1$
Well # 3	$\alpha_{opt} = 0.10$	No. of Errors = 3	$CCR = \frac{1449-3}{1449} \approx 1$	$\alpha_{opt} = 0.90$	No. of Errors = 3	$CCR = \frac{1449-3}{1449} \approx 1$
Well # 4	$\alpha_{opt} = 0.10$	No. of Errors = 1	$CCR = \frac{2041-1}{2041} \approx 1$	$\alpha_{opt} = 0.90$	No. of Errors = 1	$CCR = \frac{2041-1}{2041} \approx 1$
Well # 5	$\alpha_{opt} = 0.12$	No. of Errors = 9	$CCR = \frac{1745-9}{1745} \approx 1$	$\alpha_{opt} = 0.91$	No. of Errors = 9	$CCR = \frac{1745-9}{1745} \approx 1$
Well # 6	$\alpha_{opt} = 0.10$	No. of Errors = 3	$CCR = \frac{2845-3}{2845} \approx 1$	$\alpha_{opt} = 0.90$	No. of Errors = 3	$CCR = \frac{2845-3}{2845} \approx 1$
Well # 7	$\alpha_{opt} = 0.09$	No. of Errors = 3	$CCR = \frac{1589-3}{1589} \approx 1$	$\alpha_{opt} = 0.89$	No. of Errors = 3	$CCR = \frac{1589-3}{1589} \approx 1$
Well # 8	$\alpha_{opt} = 0.13$	No. of Errors = 2	$CCR = \frac{2271-2}{2271} \approx 1$	$\alpha_{opt} = 0.92$	No. of Errors = 2	$CCR = \frac{2271-2}{2271} \approx 1$

Table 7. Average FP plus FN for all utilized methods. Total number of data are 18,168.

Well	RF	SVM	OWA
1	78	72	14
2	51	63	4
3	36	33	3
4	14	27	1
5	35	38	9
6	37	42	3
7	24	23	3
8	18	14	2
Summation of Error	293	312	39

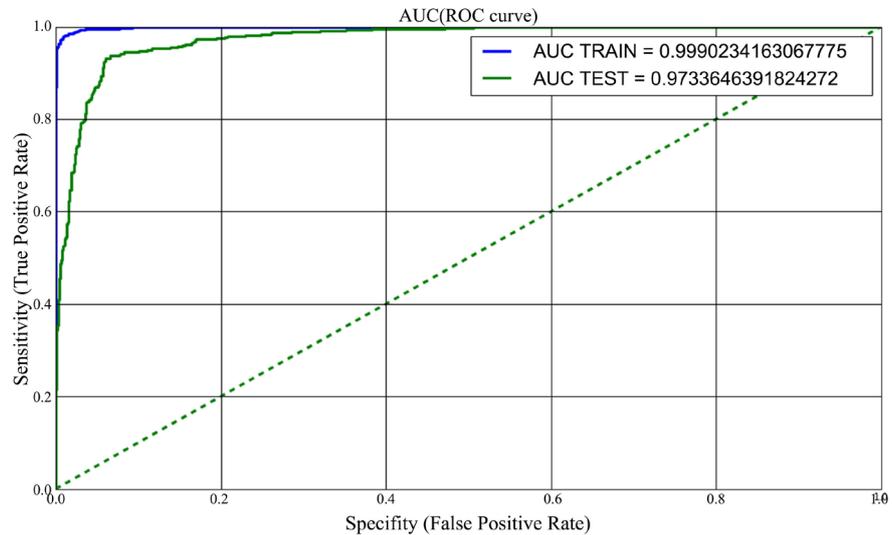


Figure 9. Receiver operating characteristic curve for SVM classifier while one well has been train and another one test.

Sensitivity is the probability of a depth will be positive given as a fracture zone. Specificity is the probability of a depth will be negative given as a nonfracture zone. The accuracy of a test is measured by the area under the ROC curve (AUC). AUC is the area between the curve and the x-axis. An area of 1 represents a perfect test, while an area of 0.5 represents a worthless test. The closer the curve follows the left-upper corner of the plot, the more accurate the test.

In **Figure 9**, the ROC curve in the case one well is used as train and another well as test, is shown. Here, SVM has been the utilized classifier. As these figures show, ROC analysis provides important information about discrimination performance: the closer the apex of the curve toward the upper left corner, the greater the discriminatory ability of the classifier. This is measured quantitatively by the AUC such that a value of >0.96 indicates excellent discriminatory ability.

6. Conclusions

In the current study, image and petrophysical logs of a carbonate-fractured reservoir of a giant oil field were studied. Therefore, conclusions depend on the utilized database, which is not necessarily general. Conclusions are briefly addressed as follows:

- DT, RF, SVM, and DL were applied to discriminate FZs from NFZs by using selected PLs in each well. The database was split into train (70%) and test (30%) and classifications were trained. Results showed that RF and SVM, with an average CCR of about 95%, give better discrimination in comparison with DT and DL. RF and SVM were selected for the generalization step, in which classifiers were trained in one of the wells and applied over the other wells. In this step, average accuracy was about 98 percent, unbelievable re-

sults.

- In the generalization step, 14 results were achieved for each well. Seven of them belong to RF classifier (7 train wells and one test well). Seven similar results were achieved while using SVM. Ordered weighted averaging was applied to integrate 14 results, to achieve one unique, more accurate, and reliable result. The accuracy for outputs for all wells was about 99%, amazing! It should be emphasized that the best accuracy in previous studies was about 72%.
- In all cases SSE for optimistic and pessimistic OWA was similar. 0.11 and 0.91 were reported for optimistic and pessimistic scenarios respectively.
- On average, the summation of the number of errors for data fusion declined from about 300 to 40. So, data fusion of more than 85 percent decreases the number of errors of classifiers.

Fracture orientation (dip and dip direction), aperture, spacing, length, filling, and filler are examples of other fracture properties that could be interpreted from ILS, and are not discussed in the current paper. Finding those properties using PLs could be the topic of future research.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] McQuillan, H. (1985) Carbonate Petroleum Reservoirs: In: Roehl, P.O. and Choquette, P.W., Eds., *Fracture-Controlled Production from the Oligo-Miocene Asmari Formation in Gachsaran and BibiHakimeh Fields, SW Iran*, Springer, New York, 511-523. https://doi.org/10.1007/978-1-4612-5040-1_33
- [2] Schlumberger (1998) Log Interpretation Principles/Applications. March, Sugar Land.
- [3] Theodoridis, S. and Koutroumbas, K. (2002) Pattern Classification. 2nd Edition, Elsevier, Athens.
- [4] Heaton, J., Goodfellow, I., Bengio, Y. and Courville, A. (2017) Deep Learning. MIT Press, Cambridge, MA.
- [5] Zhang, A., Lipton, Z.C., Li, M., Smola, A.J., Werness, B., Hu, R., Zhang, S., Tay, Y., Dagar, A. and Tang, Y. (2021) Dive into Deep Learning. <https://d2l.ai/index.html>
- [6] Yager, R.R. (1988) On Ordered Weighted Averaging Aggregation Operators in Multi Criteria Decision Making. *IEEE Transactions on Systems, Man, and Cybernetics*, **18**, 183-190. <https://doi.org/10.1109/21.87068>
- [7] Hsu, H.M. and Chen, C.T. (1996) Aggregation of Fuzzy Opinions under Group Decision Making Environment. *Fuzzy Sets and Systems*, **79**, 279-285. [https://doi.org/10.1016/0165-0114\(95\)00185-9](https://doi.org/10.1016/0165-0114(95)00185-9)
- [8] Kuncheva, L.I. and Krishnapuram, R. (1996) A Fuzzy Consensus Aggregation Operator. *Fuzzy Sets and Systems*, **79**, 347-356. [https://doi.org/10.1016/0165-0114\(95\)00148-4](https://doi.org/10.1016/0165-0114(95)00148-4)
- [9] Mazaheri, A., Memarian, H., Tokhmechi, B. and Araabi, B.N. (2015) Developing

- Fracture Measure as an Index of Fracture Impact on Well-Logs. *Energy Exploration and Exploitation*, **33**, 555-574. <https://doi.org/10.1260/0144-5987.33.4.555>
- [10] Mazahari, S.M., Memarian, H. and Tokhmechi, B. (2018) A Hybrid Learning Automata and Case-Based Reasoning for Fractured Zone Detection. *Arabian Journal of Geosciences*, **11**, Article No. 577. <https://doi.org/10.1007/s12517-018-3934-3>
- [11] Mazaheri, A., Memarian, H., Tokhmechi, B. and Araabi, B.N. (2019) Cell Size Optimization for Fracture Measure Estimation in Multi-Scale Studies within Oil Wells. *Carbonates and Evaporites*, **34**, 261-272. <https://doi.org/10.1007/s13146-017-0378-x>
- [12] Tran, N.H. (2004) Characterization and Modeling of Naturally Fractured Reservoirs [Unpublished Doctoral Dissertation]. University of New South Wales, Kennington.
- [13] Olivia Dias, L., Bom, C. R., Faria, E.L., Valentín, M.B., Correia, M.D., Albuquerque, M.P., de Albuquerque, M.P. and Coelho, J.M. (2020) Automatic Detection of Fractures and Breakouts Patterns in Acoustic Borehole Image Logs Using Fast-Region Convolutional Neural Networks. *Journal of Petroleum Science and Engineering*, **191**, 107099. <https://doi.org/10.1016/j.petrol.2020.107099>
- [14] Tokhmechi, B., Memarian, H., Rasouli, V., Noubari, H.A. and Moshiri, B. (2009) Fracture Zones Detection Using Wavelet Decomposition of Water Saturation Log. *Journal of Petroleum Science and Engineering*, **69**, 129-138. <https://doi.org/10.1016/j.petrol.2009.08.005>
- [15] Tokhmechi, B., Memarian, H., Noubari, H.A. and Moshiri, B. (2009) A Novel Approach for Fracture Zone Detection Using Petrophysical Logs. *Journal of Geophysics and Engineering*, **6**, 365-373. <https://doi.org/10.1088/1742-2132/6/4/004>
- [16] Tokhmechi, B., Memarian, H. and Rezaee, M.R. (2010) Estimation of the Fracture Density in Fractured Zones Using Petrophysical Logs. *Journal of Petroleum Science and Engineering*, **72**, 206-213. <https://doi.org/10.1016/j.petrol.2010.03.018>
- [17] Martinez-Torres, L.P. (2002) Characterization of Naturally Fractured Reservoirs from Conventional Well Logs [Unpublished Master's Thesis]. University of Oklahoma, Norman.
- [18] Jafari, A., Kadkhodaei-Ilkhchi, A., Sharghi, Y. and Ghanavati, K. (2012) Fracture Density Estimation from Petrophysical Log Data Using the Adaptive Neuro-Fuzzy Inference System. *Journal of Geophysics and Engineering*, **9**, 105-114. <https://doi.org/10.1088/1742-2132/9/1/013>
- [19] Ghoochaninejad, H.Z., Asef, M.R. and Moallemi, S.A. (2018) Estimation of Fracture Aperture from Petrophysical Logs Using Teaching-Learning-Based Optimization Algorithm into a Fuzzy Inference System. *Journal of Exploration and Production Technology*, **8**, 143-154. <https://doi.org/10.1007/s13202-017-0396-1>
- [20] Aghli, G., Soleimani, B., Moussavi-Harami, R. and Mohammadian, R. (2016) Fractured Zones Detection Using Conventional Petrophysical Logs by Differentiation Method and Its Correlation with Image Logs. *Journal of Petroleum Science and Engineering*, **142**, 152-162. <https://doi.org/10.1016/j.petrol.2016.02.002>
- [21] Moradi, M., Tokhmechi, B. and Masoudi, P. (2019) Inversion of Well Logs into Rock Types, Lithofacies and Environmental Facies, Using Pattern Recognition, a Case Study of Carbonate Sarvak Formation. *Carbonates and Evaporites*, **34**, 335-347. <https://doi.org/10.1007/s13146-017-0388-8>
- [22] Flavio, S.A., and Gregor, P.E. (1999) The Velocity-Deviation Log a Tool to Predict Pore Type and Permeability Trends in Carbonate Drill Holes from Sonic and Porosity or Density Logs. *The American Association of Petroleum Geologists Bulletin*, **83**, 450-466.
- [23] Yan, J., Lu, L., Lubbe, R. and Payne, S. (2009) Petrophysical Fracture Identification

for Rock Physics Studies. *Conference Proceedings of 71st EAGE Conference and Exhibition*, European Association of Geoscientists and Engineers, Amsterdam, 127-00193. <https://doi.org/10.3997/2214-4609.201400200>

- [24] Duda, R.O., Stork, D.G., Hart, P.E. and Hegde, N.R. (2000) *Pattern Classification*. John Wiley and Sons, Hoboken.