The Hybrid of Classification Tree and Extreme Learning Machine for Permeability Prediction in Oil Reservoir

Chandra Prasetyo Utomo

Faculty of Information Technology, Universitas YARSI
Jakarta 10510, Indonesia

Abstract

Permeability is an important parameter connected with oil reservoir. In the last two decades, artificial intelligence models have been used. The current best prediction model in permeability prediction is extreme learning machine (ELM). It produces fairly good results but a clear explanation of the model is hard to come by because it is so complex. The aim of this research is to propose a way out of this complexity through the design of a hybrid intelligent model. The model combines classification and regression. In order to handle the high range of the permeability value, a classification tree is utilized. ELM is used as a final predictor. Results demonstrate that this proposed model performs better when compared with support vector machines (SVM) and ELM in term of correlation coefficient. Moreover, the classification tree model potentially leads to better communication among petroleum engineers and has wider implications for oil reservoir management efficiency.

Keywords: *Permeability Prediction, Extreme Learning Machine, Classification Tree, Hybrid Intelligent Systems, Oil Reservoir, Regression Problem*

1. Introduction

Permeability is the flow capacity of fluid to be transmitted through a rock's pore space. According to the latest study in oil reservoir, millions of dollars can be saved or lost depending on the quality of permeability prediction. The information of permeability values in reservoirs is important because it is needed to find out the quantity of oil or gas exists in reservoirs, the quantity that can be retrieved, its flow rate, the prediction of future production, and the production facilities design. Based on that, correct knowledge of permeability is required for the whole reservoir management and development [1].

Conventional method used to obtain the permeability values is by taking rock samples in some depths then measuring its permeability in the laboratory. This method is very expensive, complex, and time consuming. In addition, laboratorial measurement is limited to the rock samples. So that, the continuous picture of permeability values can't be captured. Based on this reasons, a new method which is quite accurate, less expensive, simpler, faster, and able to deliver permeability distribution along the depth is highly needed.

A huge number of efforts have been carried out to obtain new method to predict permeability values from well log data. From 1927 to 1981, scientists had tried empirical models by delivering mathematical formulas to get permeability values. None of this formula gives satisfying result in general case. Since 1961, multiple variable regressions models had been applied. The distribution of predicted values gained from this model is still far from actual values. However, empirical and regression models gave hint about factors controlling permeability [2].

In the past two decades, computational intelligent techniques, such as artificial neural networks (ANN), have been utilized in permeability prediction. An ANN is a powerful and flexible tool for many applications including in petroleum area. This model is able to learn from previous data in order to predict values from new data. It gives better performance than previous models in predicting permeability from well logs in new wells [3]. Nevertheless, back propagation neural network suffers some drawbacks. It has some tuning parameters such as number of hidden neurons, learning rate, and momentum so it needs more efforts to find the best model. In addition, the gradient based learning algorithm used by ANN makes the training process becomes time consuming.

Many works have been tried to develop new ANN model to solve its weaknesses. In 2004, Huang [4] proposed new learning algorithm for single-hidden layer feed forward neural networks which is called extreme learning machine (ELM). Both in theory and experimental results, this learning algorithm gives better generalization performances and extremely faster learning speed than traditional popular gradient based learning algorithm [5]. Based on that, ELM has been highly exploited in many applications including in petroleum engineering area. In comparison with support vector machines (SVM) and conventional ANN for predicting permeability from well log data, ELM gives better generalization ability and faster speed [1]. This result stated that ELM is the current best single model in permeability prediction problem.

Although ELM gives fairly good results and faster speed, it still has some limitations. First, ELM can't deal with high data distribution of permeability values. One of the main challenges in predicting permeability is high range of its values in each well [6]. Second, ELM can't give knowledge representation of developed model. Because of its structure which is dense combination of simple computation, trained ELM is hard and complex to be written in mathematical formulas. As a result, it is impossible to produce understandable knowledge representation which is needed to communicate with expert for future study and research.

In this research, a new hybrid intelligent model which can manage high data distribution and give knowledge representation is proposed. To deal with high range data, a single model is not enough. The data should be classified into low permeability and high permeability then applied different models to predict the value.

This proposed hybrid model is basically combination of classification and regression models. Classification model is responsible to classify the data into low and high permeability. On the other hand, regression models are responsible to give final prediction value of its associated data. Classification tree is utilized as classification model since it can produce knowledge representation which is close to human intuition. ELM is used as regression model since it is currently the best single model in permeability prediction.

The rest of this paper is organized as the following. Section 2 is dedicated as previous works. In this section, review in permeability prediction and overview of ELM are presented. In Section 3, design of the proposed model and its implementation are explained. In Section 4, experiments, results, and analysis are provided. Finally, conclusions and future works are given in Section 5.

2. Previous Works

There are huge efforts from scientists and engineers in order to deliver best model to predict permeability values based on well logs data. This section describes previous works in permeability prediction which can be categorized into empirical models, multiple regression variable models, and artificial intelligence models.

2.1 Empirical Models

Empirical models are predicting permeability by defining mathematical formulas based on its correlation with some rock properties. Kozeny [2] introduced the first equation of permeability in 1927. He measured permeability as a function of empirical Kozeny constant, porosity, and surface area. Archie [7] established the concept of "formation resistivity factor" in 1941. His concept indirectly influenced the computation of permeability since it affected the way to calculate water saturation.

Tixier [8] proposed a formula in 1949 to determine permeability from resistivity gradients by using empirical correlation between resistivity and water saturation, water saturation and capillarity pressure, and capillarity pressure and permeability. In 1950, Wyllie & Rose [9] modified the formula proposed by Tixier. Their model is based on quantitative log interpretation theoretical analysis and some assumptions.

In 1956, Sheffield [10] delivered permeability formula based on Kozeny's equation and formation of a correlation coefficient for some water well-known water-wet sands. However, he recommended his formula is suitable only for clean sands. In 1963, Prison [10] proposed formula which was determined by multiple correlation from relatively few data. For high gravity crudes (API > 40o) and for depths greater than 6500 ft, the formula must not be utilized.

Timur [11] generalized permeability equation based on the work of Kozeny and Willy & Rose. In 1974, Coates & Dumanoir [12] proposed an improved empirical permeability formula which is satisfied the condition of zero permeability at zero porosity and when irreducible water saturation is 100%. Coates and Denoo [13] simplified the previous proposed formulas and still satisfied the zero permeability condition. However, the formation must be at irreducible water saturation.

2.2 Multiple Variable Regression Models

Multiple variable regression models are expansions of the regression analysis that include extra independent variables in the equation. The model can be generalized as:

$$Y = C_0 + C_1X_1 + C_2X_2 + \cdots + C_nX_n + e \qquad (1)$$

where Y is the dependent variable, X_1, X_2,…, X_n are the independent variables, and e is a random error or residual. The regression coefficients C_1, C_2,…,C_n are the parameters to be approximated.

A general procedure of multiple variable regression for permeability prediction was established by Wendt and Sakurai [14] in 1986. The main drawback of using this model is the predicted permeability values is narrower than the actual values. Kendall and Stuart [15] enlightened above phenomena by stating this model gave the best prediction on the average. Weighting the high and low values are applied to improve the capability of regression model to predict outlier data. However, this may turn the predictor into unstable and statistically biased. Pereira [16] reported that density, derivative of density, gamma ray, and derivative of gamma ray are the best combination to be utilized as independent variables in multiple regression analyses.

2.3 Artificial Intelligence Models

Artificial Intelligence (AI) is set of models inspired by nature such as neural networks, fuzzy logic, and genetic algorithm. A lot of neural networks applications can be found in the petroleum industry, from exploration, drilling exploration, to reservoir and production engineering [17]. In predicting permeability, neural networks gave significant improvement [18-21]. This opened the door of others AI models to be applied in the petroleum industry area especially in the permeability prediction problem.

The combination of two or more AI models is called hybrid model. It complements the weaknesses of one model with the advantage of others. Since neural networks is one of the best AI model, most of published hybrid model are neural network based model. There are some proposed hybrid models in permeability prediction. Deni [22] proposed a hybrid of genetic algorithm and fuzzy/neural network inference system. Helmi [23] developed a hybrid of fuzzy logic, support vector machine, and functional network. Karimpouli [6] built up supervise committee machine neural network. Li [24] enhanced decision tree learning approach for neural decision tree model.

Although previous hybrid model gave better results than single model, it has some drawbacks due to the limitation of neural networks model. As a "black box" model, neural networks cannot give clear relationships among variables. Other limitations are it can fall into local minima, need to adjust too many parameters, and time consuming.

2.4 Extreme Learning Machines

A lot of works has been tried to resolve the drawbacks of ANN. Huang and Babri [25] proved that single hidden layer feedforward neural networks (SLFN) with at most *m* hidden nodes is able to approximate function for *m* distinct vectors in training dataset.

Let given *m* vectors in training dataset $\mathbf{D} = \{(\mathbf{x}^{(k)}, \mathbf{t}^{(k)}) \mid \mathbf{x}^{(k)} \in \mathbf{R}^n, \mathbf{t}^{(k)} \in \mathbf{R}^p, k = 1,..,m\}$ where $\mathbf{x}^{(k)} = [x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}]^T$ and $\mathbf{t}^{(k)} = [t_1^{(k)}, t_2^{(k)}, \ldots, t_p^{(k)}]^T$. A SLFN with M hidden nodes, activation function $g(x)$ in hidden nodes, and linear activation function in output nodes is mathematically modeled as:

$$\sum_{i=1}^{M} \beta_i g_i(\mathbf{x}^{(k)}) = \sum_{i=1}^{M} \beta_i g(\mathbf{w}_i \cdot \mathbf{x}^{(k)} + b_i) = \mathbf{o}^{(k)}, \quad k = 1, \ldots, m \tag{2}$$

where

$\mathbf{w}_i \in \mathbf{R}^n$ is the weights attached to the edge connecting input nodes and the *i*-th hidden node

$$\mathbf{w}_i = [w_{i1}, w_{i1}, \ldots, w_{in}]^T, \tag{3}$$

$\boldsymbol{\beta}_i \in \mathbf{R}^p$ is the weights attached to the edge connecting the *i*-th hidden node and the output nodes

$$\boldsymbol{\beta}_i = [\beta_{i1}, \beta_{i2}, \ldots, \beta_{ip}]^T, \tag{4}$$

$\mathbf{w}_i \cdot \mathbf{x}^{(k)}$ is the inner product of $\mathbf{w}_i$ and $\mathbf{x}^{(k)}$,
b_i is the bias of the *i*-th hidden node,
$\mathbf{o}^{(k)} \in \mathbf{R}^p$ is the output of neural network for *k*-th vector.

The meaning of SLFN can approximate *m* vectors is there are exist $\mathbf{w}_i$, $\boldsymbol{\beta}_i$, and b_i, such that:

$$\| \mathbf{o}^{(k)} - \mathbf{t}^{(k)} \| = 0 \tag{5}$$

$$\sum_{i=1}^{M} \beta_i g(\mathbf{w}_i \cdot \mathbf{x}^{(k)} + b_i) = \mathbf{t}^{(k)}, \quad k = 1, \ldots, m \tag{6}$$

Those *m* equations can be written as:

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{T}, \tag{7}$$

where

$\mathbf{H} \in \mathbf{R}^{m \times M}$ is the hidden layer output matrix of the neural networks.

$$\mathbf{H} = \begin{bmatrix} g(w_1 \bullet x^{(1)} + b_1) & \cdots & g(w_M \bullet x^{(1)} + b_M) \\ \vdots & \ddots & \vdots \\ g(w_1 \bullet x^{(m)} + b_1) & \cdots & g(w_M \bullet x^{(m)} + b_M) \end{bmatrix} \tag{8}$$

$\boldsymbol{\beta} \in \mathbf{R}^{M \times p}$ is the weights connecting hidden layer and output layers

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1{}^T \\ \vdots \\ \beta_M{}^T \end{bmatrix}, \tag{9}$$

$\mathbf{T} \in \mathbf{R}^{m \times p}$ is the target values of *m* vectors in training dataset

$$\mathbf{T} = \begin{bmatrix} t^{(1)T} \\ \vdots \\ t^{(m)T} \end{bmatrix}, \tag{10}$$

In the conventional gradient descent based learning algorithm, weights $\mathbf{w}_i$ which is connecting the input layer and hidden layer and biases b_i in the hidden nodes are needed to be initialized and tuned in every iteration. This is the main factor which often makes training process of neural networks become time consuming and the trained model may not reach global minima.

Huang [5] proposed minimum norm least-squares solution of SLFN which doesn't need to tune those parameters. Training SLFN with fixed input weights $\mathbf{w}_i$ and the hidden layer biases b_i is similar to find a least square solution $\widehat{\boldsymbol{\beta}}$ of the linear system $\mathbf{H}\boldsymbol{\beta} = \mathbf{T}$:

$$\left\| \mathbf{H}(\mathbf{w}_1, \ldots, \mathbf{w}_M, b_1, \ldots, b_M,)\widehat{\boldsymbol{\beta}} - \mathbf{T} \right\| = \min_{\boldsymbol{\beta}} \left\| \mathbf{H}(\mathbf{w}_1, \ldots, \mathbf{w}_M, b_1, \ldots, b_M,)\boldsymbol{\beta} - \mathbf{T} \right\|. \tag{11}$$

The smallest norm least squares solution of the above linear system is

$$\widehat{\boldsymbol{\beta}} = \mathbf{H}^{\dagger}\mathbf{T} \tag{12}$$

where $\mathbf{H}^{\dagger}$ is the *Moore-Penrose generalized inverse* of matrix H. This solution has three important properties which are minimum training error, smallest norm of weights, and unique solution which is $\widehat{\boldsymbol{\beta}} = \mathbf{H}^{\dagger}\mathbf{T}$.

The above minimum norm least-square solution for SLFN is called extreme learning machine (ELM). Let given *m* vectors in training dataset $\mathbf{D} = \{(\mathbf{x}^{(k)}, \mathbf{t}^{(k)}) \mid \mathbf{x}^{(k)} \in \mathbf{R}^n, \mathbf{t}^{(k)} \in \mathbf{R}^p, k = 1,..,m\}$, activation function $g(x)$, and number of hidden node M. The training process of ELM is the the following:

Step (1) Randomly set input-hidden layer weights $\mathbf{w}_i$ and bias b_i, $i = 1,\ldots,M$.

Step (2) Compute the matrix of hidden layer output $\mathbf{H}$

Step (3) Compute the hidden-output layer weights $\widehat{\boldsymbol{\beta}}$ for $\widehat{\boldsymbol{\beta}} = \mathbf{H}^{\dagger}\mathbf{T}$ where $\mathbf{T} = [\mathbf{t}^{(1)},\ldots, \mathbf{t}^{(m)}]$.

The comparison between conventional widely used neural networks and ELM is summarized in the Table 1.

Table 1: The comparison between Back Propagation ANN and ELM

No.	Points of Comparison	Comparison
1.	Learning Algorithm	ANN: Gradient based learning ELM: Minimum least-squares
2.	Training Parameters	ANN: Need to tuning number of hidden nodes, learning rate, momentum, and termination criteria ELM: Simple tuning-free algorithm. The only one to be defined is number of hidden nodes
3.	Activation Function	ANN: Works only for differentiable functions ELM: Works for differentiable and many non-differentiable functions
4.	Speed	ANN: Very slow especially in the large dataset. All of weights are updated in every iteration. ELM: Extremely faster than BP ANN. Only three steps without any iteration
5.	Result	ANN: Get trained model which has minimum training error. There is possibility to finish in the local minima. ELM: Get trained model which has minimum training error and smallest norm of weight. Better generalization model and reach global minima.

IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 1, No 1, January 2013
ISSN (Print): 1694-0784 | ISSN (Online): 1694-0814
www.IJCSI.org

3. Design and Implementation Model

The main challenge in permeability prediction is high range of permeability. A single model is not enough to deal with that. The data should be classified into low permeability and high permeability then applied different model to predict the value.

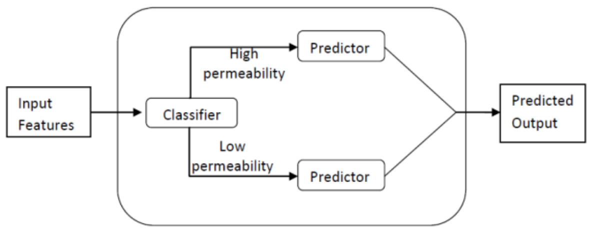

Fig. 1 Design of the proposed hybrid model

Hybrid model which is basically combination of classification and regression models is proposed. Classification model is responsible to classify the data into low and high permeability based on a threshold value. On the other hand, regression models are responsible to give final prediction value of its associated data. Design of this model can be seen in the Fig 1.

One of the objectives in this research is to propose new model which gives understandable knowledge representation. The best representation model which is close to human reasoning is classification tree. For this reason, Classification Tree model is used in the classification part. Since Classification and Regression Tree (CART) from Salford System [26] is one of the best tools for classification tree design, it is implemented in this proposed model.

As presented in previous section, ELM is the current best single model in permeability prediction. ELM developed by Huang [27] is implemented in this proposed model as final predictor.

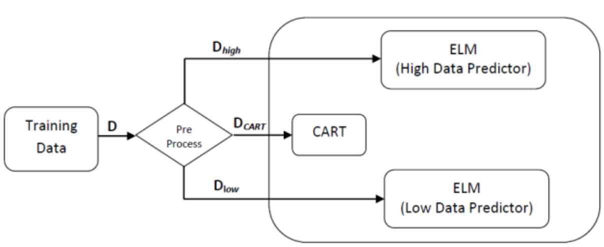

Fig. 2 Training procedure of proposed hybrid model

Let we have m vectors in training dataset $\mathbf{D}$.

$$\mathbf{D} = \{(\mathbf{x}^{(k)}, t^{(k)}) \mid \mathbf{x}^{(k)} \in \mathbf{R}^n, k = 1,..,m.\}. \qquad (13)$$

The training algorithm of this hybrid model is designed as the following:

Step (1) Add Discretized Target

Discretize the target output $t^{(k)}$ into two classes *"low"* and *"high"* based on selected threshold value. The new training dataset is $\mathbf{D}_1 = \{(\mathbf{x}^{(k)}, t^{(k)}, t_d^{(k)}) \mid \mathbf{x}^{(k)} \in \mathbf{R}^n, k = 1,..,m.\}$ with $t_d^{(k)}$ is "low" if $t^{(k)} \le$ threshold, otherwise $t_d^{(k)}$ is "high".

Step (2) Produce the Associated Training Data

In this step, three training dataset $\mathbf{D}_{CART}$, $\mathbf{D}_{low}$, $\mathbf{D}_{high}$ are produced. The training dataset for CART $\mathbf{D}_{CART}$ is $\mathbf{D}_1$ without original target value $t^{(k)}$. The vector $(\mathbf{x}^{(k)}, t^{(k)}, t_d^{(k)})$ in $\mathbf{D}_1$ is putted into $\mathbf{D}_{low}$ if $t_d^{(k)}$ = "low", otherwise it is putted into $\mathbf{D}_{high}$. The $t_d^{(k)}$ element in the $\mathbf{D}_{low}$ and $\mathbf{D}_{high}$ are removed at the end of this step.

Step (3)Train the CART

Train the CART by training dataset

$$\mathbf{D}_{CART} = \{(\mathbf{x}^{(k)}, t_d^{(k)}) \mid \mathbf{x}^{(k)} \in \mathbf{R}^n, k = 1,..,m.\} \qquad (14)$$

Step (4) Train the ELMs

Train the low ELM by training dataset

$$\mathbf{D}_{low} \{(\mathbf{x}^{(l)}, t^{(l)}) \mid \mathbf{x}^{(l)} \in \mathbf{R}^n, l = 1,..,y.\} \qquad (15)$$

Train the high ELM by training dataset

$$\mathbf{D}_{high} \{(\mathbf{x}^{(h)}, t^{(h)}) \mid \mathbf{x}^{(h)} \in \mathbf{R}^n, h = 1,..,z.\} \qquad (16)$$

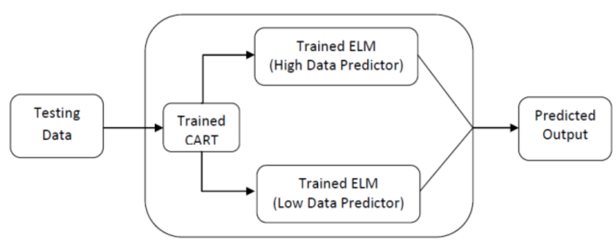

Fig. 3 Testing procedure of trained model

After finish four steps above, the trained hybrid model is produced and ready to predict permeability from new dataset. The illustration can be seen in Fig 3.

IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 1, No 1, January 2013
ISSN (Print): 1694-0784 | ISSN (Online): 1694-0814
www.IJCSI.org

4. Experiments, Results, and Analysis

The data used in this experiment are 5 well logs data from Saudi Aramco. Data for Well 1 has 145 rows (vectors), for Well 2 has 141 rows, for Well 3 has 193 rows, for Well 4 has 147 rows, and for Well 5 has 141 rows. There are 5 input variables which are DT (sonic travel time), GR (Gamma Ray), PHIE (Effective Porosity), RHOB (Density), and SWT (Water Saturation). The target output to be predicted is PERM (Permeability).

Two kinds of experiments are conducted in this research. In the first experiment, one well is chosen as tested well and the rest wells are used to train the model. Because there are 5 wells, this experiment is repeated up to 5 times with different combination of training and testing wells. In the second experiment, all data are combined then divided randomly into training and testing data with ratio 80:20. The training data is used to train the model. Then, the trained model is tested by testing data to predict the permeability values.

The input features are normalized into [-1,1] and the output target is kept in the original value. The threshold used in this experiment is 1. This means, if the permeability value is less or equal than 1, then it is considered as low permeability. Otherwise, it is high permeability. A number of experiments had been tried to get the best parameters combination of CART such as in splitting criteria, stopping conditions, and thresholds.

Both classification and final prediction performance will be measured. The performance measurements for classification are Accuracy (ACC), True Positive Rate (TPR), and False Positive Rate (FPR). In order to measure the performance of the whole model, Root Mean Square Error (RMSE) and Correlation Coefficient (R) are used as performance criteria. The proposed model will be compared with SVM [28] and ELM based on this performance criteria.

ELM assigns randomly input weights and biases in the first step of execution. To reduce the influence of random generator, 10 sequences of executions are applied in each model and the average results are obtained.

Table 2: The performances of CART as classifier

Tested Well	TPR	FPR	ACC
1	0.8333	0.4220	0.6414
2	0.8333	0.0693	0.9148
3	0.5783	0.2818	0.6500
4	0.2727	0.0326	0.7075
5	0.4658	0.0735	0.6879

Table 3: The performances comparison of models

Tested Well	RMSE			R		
	SVM	ELM	Hybrid	SVM	ELM	Hybrid
1	7.87	9.77	12.24	0.55	0.44	0.44
2	19.47	14.48	13.98	0.67	0.77	0.73
3	16.82	15.52	15.29	0.38	0.39	0.42
4	9.38	8.514	9.51	0.40	0.44	0.35
5	10.40	8.42	9.60	0.38	0.47	0.47

The performances of CART as classifier to classify the high and low permeability data are shown in Table 2. These performances are obtained after tree pruning. When there is no pruning mechanism in classification tree induction, the classifier testing performances are bad and the final predictions of hybrid model are not reliable. Table 3 shows that the performances of proposed model are similar with current single best prediction model in permeability prediction.

The comparison of models based on RMSE can be clearly seen in Fig. 4. Except the models for tested Well 1, SVM models give the highest errors. The proposed models are better than ELMs in tested Wells 2 and Well 3.

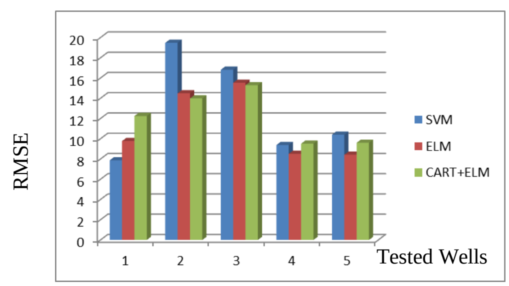

Fig. 4 The performances comparison based on RMSE

Fig. 5 shows the comparison of models based on Correlation Coefficient R. SVMs give the worst performances in tested Wells 2, 3, and 5. The proposed models are better than ELMs in tested Well 3 and 5, worse in tested Wells 2 and 4, and almost equal in tested Well 1.

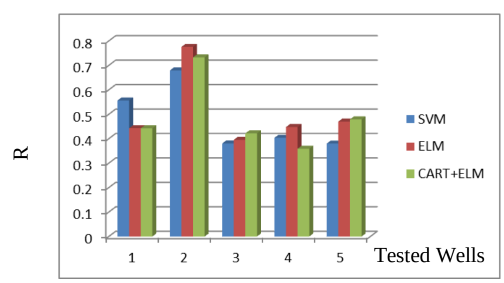

Fig. 5 The performances comparison based on Corr. Coefficient (R)

The performances results of the second experiment which is randomly divided data into training and testing data can be seen in table 4. In term of RMSE, the proposed model is worse than SVM and ELM. In term of R, the proposed model is better than SVM and ELM.

Table 4: The performances comparison of models in general Wells

Model	RMSE	R
SVM	12.88770	0.20670
ELM	12.49098	0.23453
CART + ELM	13.07807	0.26898

Another way to see differences of prediction is by looking the plot of actual and predicted values. Fig. 6 gives the permeability data plot of actual value and predicted value by ELM and proposed model. This figure shows that the proposed model can handle high distribution data and predict accurately the low permeability values. However, it is still not good enough to predict the high permeability values.

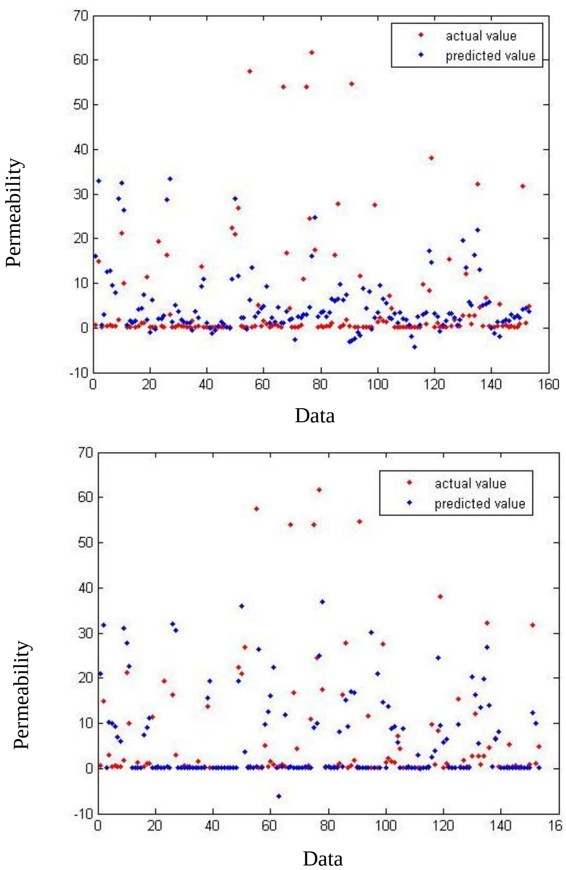

Fig. 6 Plotting permeability data of actual values and predicted values by ELM (top) and CART+ELM (bottom)

One of the most important objectives in this research is deliver knowledge representation. The classification tree is produced in the classification part. The classification tree produced in the second experiment can be seen in the Fig.7. This tree is simple and understandable. Some rules connected with relationship between permeability and the predictors can be drawn. It can be used to communicate with experts and researchers in domain problem.

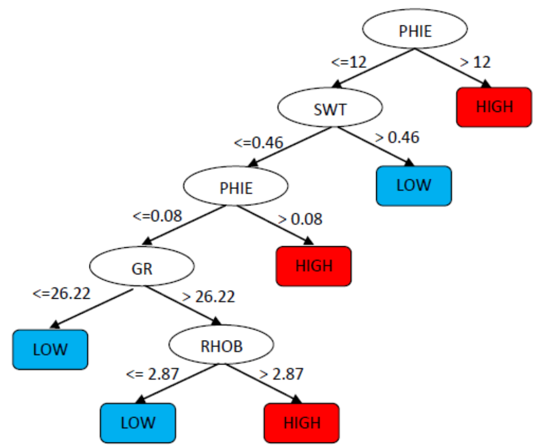

Fig. 7 Classification tree generated by CART in the classification part

5. Conclusions and Future Works

Based on the results and analysis of the experiments, some conclusions can be drawn. The proposed hybrid model, which is combination of Classification Tree as classifier and ELM as predictor, gives better performance than SVM and ELM in term of correlation coefficient in general Wells. The prediction in low permeability data is excellent but still not good enough in high permeability data.

The classification part plays important role in determining the prediction. The better accuracy of classifier, the better result in final prediction. The classification tree produced by this hybrid model is simple and understandable. This means, it will be promising tool to be widely used to communicate with domain expert. Although the proposed model just gave small improvement, it concludes that the use of hybrid model in this way is in the right direction.

The future work will be improvement in both classification and regression parts of this hybrid model. It is interesting to see how performance of classification tree with others induction tree algorithms. It is also necessary to investigate different possible hybrid models which combine classification tree with other regression models such as support vector regressions and fuzzy systems.

Acknowledgments

This paper is result from master thesis works by the author at the King Abdullah University of Science and Technology (KAUST). He would like to thank Prof. Mikhail Moshkov for supports and guidance. He would also like to thank Dr. Igor Chikalov for technical supports, Dr. Xiangliang Zhang for insight of using classification, and Dr. L. Ghouti for providing well logs data from Saudi Aramco Oil Reservoir. Many thanks also to Prof. Basem Shihada and Prof. Shuyu Sun for examining and approving this work as thesis committee member.

References

[1] S. O. Olatunji, *et al.*, "Modeling Permeability Prediction Using Extreme Learning Machines," in *Mathematical/Analytical Modelling and Computer Simulation (AMS), 2010 Fourth Asia International Conference on*, 2010, pp. 29-33.

[2] B. Balan, *et al.*, "State-Of-The-Art in Permeability Determination From Well Log Data: Part 1- A Comparative Study, Model Development," in *SPE Eastern Regional Conference & Exhibition*, Morgantown, West Virginia, USA, 1995.

[3] S. Mohaghegh, *et al.*, "State-Of-The-Art in Permeability Determination From Well Log Data: Part2- Verifiable, Accurate Permeability Prediction, The Touch-Stone of All Models," in *SPE Eastern Regional Conference & Exhibition*, Morgatown, West Virginia, USA, 1995.

[4] H. Guang-Bin, *et al.*, "Extreme learning machine: a new learning scheme of feedforward neural networks," in *Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on*, 2004, pp. 985-990 vol.2.

[5] G.-B. Huang, *et al.*, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, pp. 489-501, 2006.

[6] S. Karimpouli, *et al.*, "A new approach to improve neural networks' algorithm in permeability prediction of petroleum reservoirs using supervised committee machine neural network (SCMNN)," *Journal of Petroleum Science and Engineering*, vol. 73, pp. 227-232, 2010.

[7] G. E. Archie, "The Electrical Resistivity Log as an Aid in Determining Some Reservoir Charachteristics," *Trans., AIME*, vol. 146, pp. 54-62, 1942.

[8] M. P. Tixier, "Evaluation of Permeability From Electric-Log Resistivity Gradients," *Oil and Gas Journal*, p. 113, June 1949.

[9] M. R. J. Wyllie and W. D. Rose, "Some Theoretical Considerations Related to the Quantitative Evaluation of the Physical Characteristics of Reservoir Rock from Electric Log Data," *Trans., AIME*, vol. 189, p. 105, 1950.

[10] S. J. Pirson, *Handbook of Well Log Analysis*: Englewood Cliffs, N.J., Prentice-Hall, Inc., 1963.

[11] A. Timur, "An Investigation of Permeability, Porosity, and Residual Water Saturation Relationship for Sandstone Reservoirs," *The Log Analyst*, vol. 9, p. 8, July-August 1968.

[12] G. R. Coats and J. L. Dumanoir, "A New Approach to Improved Log-Derived Permeability," *The Log Analyst*, January-February 1974.

[13] SchlumbergerLtd., *Log Interpretation Charts*. Houston, Texas, 1987.

[14] W. A. Wendt, *et al.*, Eds., *Permeability Prediction From Well Logs Using Multiple Regression* (Reservoir Characterization. New York: Academic Press, 1986.

[15] N. R. Draper and H. Smith, *Applied Regression Analysis*: Wiley, 1981.

[16] J. L. L. Pereira, "Permeability prediction from well log data using multiple regression analysis," M.S.PNGE. 1424328, West Virginia University, United States -- West Virginia, 2004.

[17] M. P. McCormak, "Neural Networks in the Petroleum Industry," in *Society of Exploration Geophysicists 1991 Technical Program*, 1991, pp. 285-289.

[18] S. Mohaghegh, *et al.*, "A Methodological Approach for Reservoir Heterogeneity Characterization Using Artificial Neural Networks," presented at the SPE Annual Technical Conference, New Orleans, LA, September 1994.

[19] S. Mohaghegh, *et al.*, "Design and Development of An Artificial Neural Network for Estimation of Formation Permeability," presented at the SPE Petroleum Computer Conference, Dallas, Texas, August 1994.

[20] D. A. Osborne, "Permeability Estimation Using a Neural Network: A Case Study from The Roberts Unit, Wasson Field, Yoakum County, Texas," *AAPG South West Section Transactions*, pp. 125-132, 1992.

[21] J. M. Wiener, "Predicting Carbonate Permeabilities from Wireline Logs Using a Back-Propagation Neural Network," in *Society of Exploration Geophysicists 1991 Technical Program*, 1991.

[22] X. Deny, *et al.*, "Permeability Estimation Using Hybrid Genetic Programming and Fuzzy/Neural Inference Approach," presented at the Society of Petroleum Engineers Annual Technical Conference and Exhibition, Dallas, Texas, USA, 9-12 October 2005.

[23] T. Helmy, *et al.*, "Hybrid computational models for the characterization of oil and gas reservoirs," *Expert Systems with Applications*, vol. 37, pp. 5353-5363, 2010.

[24] X. Li and C. W. Chan, "Application of an enhanced decision tree learning approach for prediction of petroleum production," *Engineering Applications of Artificial Intelligence*, vol. 23, pp. 102-109, 2010.

[25] H. Guang-Bin and H. A. Babri, "Upper bounds on the number of hidden neurons in feedforward networks with arbitrary bounded nonlinear activation functions," *Neural Networks, IEEE Transactions on*, vol. 9, pp. 224-229, 1998.

[26] S. System, "Salford Predictive Modeler Builder v6.6," ed. San Diego, California: Salford System, 2011.

[27] G.-B. Huang, "ELM Source Codes: http://www.ntu.edu.sg/home/egbhuang/," ed, 2011.

[28] C.-C. Chang and C.-J. Lin, " LIBSVM : a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1--27:27, 2011.

C. P. Utomo received B.Sc in Computer Science from University of Indonesia and M.S in Computer Science from King Abdullah University of Science and Technology (KAUST) in 2009 and 2011, respectively. Currently, he is researcher and lecturer at the Faculty of Information Technology Universitas YARSI, Jakarta. His research interests are artificial intelligence, machine learning, and data mining. Specifically, he is eager in applying computational intelligent methods such as neural networks, extreme learning machine, and support vector machines in real world problems in the specific domain such as classification, regression, association rules, and clustering.