



## Lithological facies classification using deep convolutional neural network

Yadigar Imamverdiyev, Lyudmila Sukhostat\*

Institute of Information Technology, Azerbaijan National Academy of Sciences, 9A, B. Vahabzade Street, AZ1141, Baku, Azerbaijan



### ARTICLE INFO

#### Keywords:

Well log data  
Facies  
Facies classification  
Deep learning  
Convolutional neural network

### ABSTRACT

The aim of this paper is the development of an effective model based on deep learning for geological facies classification in wells. Facies classification is carried out by studying the lithological properties of rocks, which are characteristic of modern sediments, accumulating in certain physical and geographical conditions. In this study, a new 1D-CNN model, which is trained on various optimization algorithms, is proposed. The photoelectric effect, gamma ray, resistivity logging, neutron-density porosity difference, average neutron density porosity and geologic constraining variables are considered as input data of the model. Acceptable accuracy and the use of conventional well log data are the main advantages of the proposed intellectual model. The proposed model is compared with a recurrent neural network model, a long short-term memory model, a support vector machine model, and a k-nearest neighbor model and shows more accurate results in comparison with them. The model shows successful results in the study of well log data and can, therefore, be recommended as a suitable and effective approach for well log data processing required for lithological discrimination.

### 1. Introduction

The most popular geophysical research of the well, conducted with the purpose of revealing oil in a geological section, is logging. Well logs are physical and chemical measurements of rocks recorded by lowering specialized sensors into the wells after drilling. It allows determining the porosity, permeability, fluid composition, information about oil and gas saturation, etc. With the help of logging the character of the drilled layers is determined without core selection.

The facies (part of the layer that differs from the adjacent layers by lithological composition) contain characteristic features found in the core of reservoir rock samples that were taken from wells. They are the basis for characterizing the reservoir and constructing its model. Experts analyze them to determine the type and sequence of facies. This task is tedious and time-consuming. The process of wells drilling and cores obtaining for well analysis is expensive. These costs increase significantly as the number of wells increases. Therefore, it is necessary to develop approaches for facies determination using well log data.

Oil wells' characteristics are measured by various petrophysical instruments that help petrophysicists to differentiate the facies. The problem of facies prediction based on well log data was solved using various approaches.

Most researchers suggest using artificial neural networks (ANN). Neural networks can solve problems that cannot be solved with the help of conventional calculations and discover very complex relationships

between several variables. ANN methods have a remarkable ability to establish a complex mapping between nonlinearly coupled input and output data (Nakutnyy et al., 2008). In petroleum engineering, these networks are used when there is not enough data to interpret (Auda and Kamel, 1999) (Ayala and Ertekin, 2005).

ANN is used to estimate all unknown reservoir parameters. The coefficients of interpolating Chebyshev polynomials were considered as input to the ANN (Adibifard et al., 2014). Different training algorithms used to train ANN, and the optimum number of neurons for each algorithm were obtained by minimizing the mean relative error (MRE) over test data. Levenberg-Marquardt algorithm showed the best result.

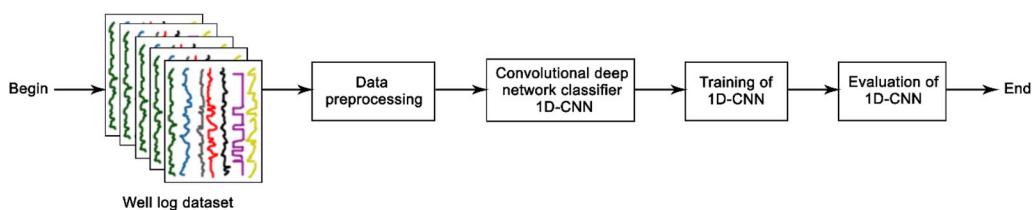
Methods based on two different types of intelligent approaches, including ANN linked to the particle swarm optimization (PSO) tool, was developed to evaluate the productivity of horizontal oil wells (Ahmadi et al., 2015). The authors of the paper suggest that the presented prognostic model can be used for effective forecasting of well productivity, in particular, at the initial stages of the evolvement of horizontal well drilling.

An approach based on the principal component analysis (PCA) and ANN was proposed (Gao et al., 2016) to implement an accurate and effective reservoir well productivity prediction using fluvial facies. The proposed approach takes into account the reservoir complexity and the filtering mechanisms to predict the well productivity.

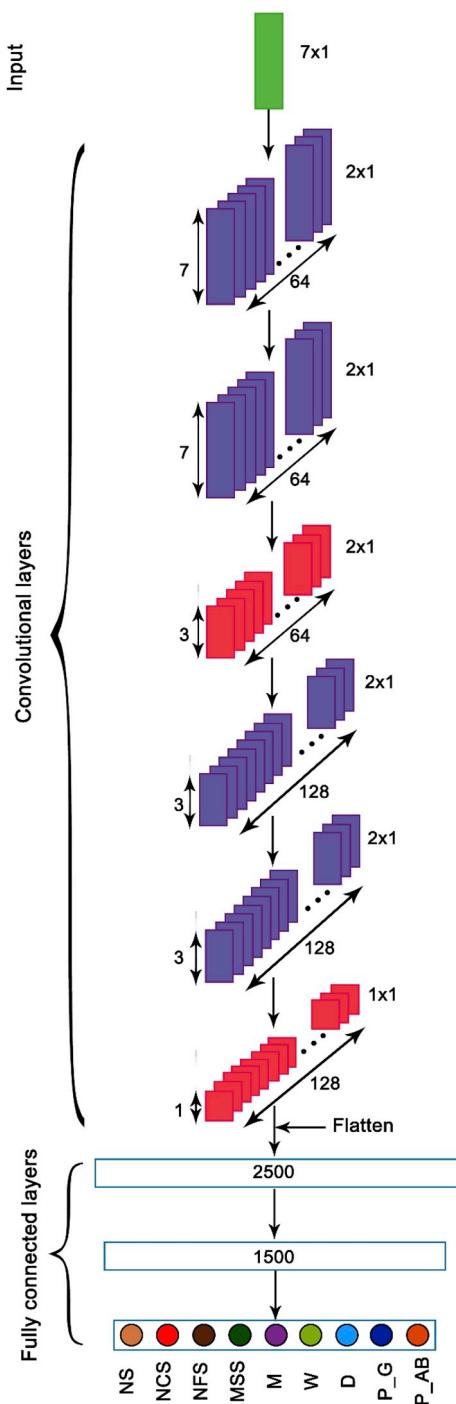
Porosity was noted as the main attractive attribute, and selection of petrofacies is the preferred way of permeability estimation in the

\* Corresponding author.

E-mail addresses: [yadigar@lan.ab.az](mailto:yadigar@lan.ab.az) (Y. Imamverdiyev), [lsuhostat@hotmail.com](mailto:lsuhostat@hotmail.com) (L. Sukhostat).



**Fig. 1.** Flowchart of the proposed approach.



**Fig. 2.** The overall architecture of 1D-CNN.

uncored wells (Chehrazi and Rezaee, 2012). The fuzzy c-means (FCM) clustering method was applied for the subdivision of the data space into petrofacies, and the corresponding relationships between porosity and permeability for each facies were determined.

An alternative method of porosity prediction, which is based on integration between wavelet theory and ANN, was presented (Saljooghi and Hezarkhani, 2014). In this study, different wavelets are applied as activation functions to predict the porosity from well log data.

Data mining and machine learning approach with MICP (mercury injection capillary pressure) was proposed for predicting the pore structure of Mesozoic strata (Wang et al., 2018). Pore structure parameters were calculated and characterized according to MICP data. The paper explains the capability of well log and core data reflect the information of pore structure. The model is established by the decision tree method with the overall accuracy rate of 90.9% and the cross-validation rate 75%.

An approach based on the ANN, adaptive neuro-fuzzy inference system (ANFIS), and support vector machine (SVM) was developed (Elkatatny and Mahmoud, 2018). The approach was designed to determine the oil formation volume factor (OFVF) based on the specific gravity of gas, the dissolved gas to oil ratio, the oil specific gravity, and the temperature of the reservoir.

Log data is used for lithological analysis, which is an integral part of facies analysis, taking into account lithological features of rocks (composition, structure, presence of mineral indicators of the environment, etc.) (Serra, 1984). Rocks that have numerous physical and chemical properties can be used in classification.

Self-organizing map neural network (SOM-ANN) and hierarchical cluster analysis (HCA) were utilized to characterize lithofacies in uncored but logged wells (Sifidari et al., 2014). The electrofacies derived from the SOM networks showed a good agreement with reservoir geological (lithofacies) and petrophysical data.

Variations in petrophysical lithofacies were evaluated, and structural facies-controls were identified (Ohl and Raef, 2014). A neural network petrophysical facies classification was based on training and validation using three petrophysically-different wells and three volume seismic attributes, including the wavelet of the reservoir-top reflection.

Natural fractures have a significant influence on the petroleum reservoir behavior and performance. Three types of fractures are observed: opened, sealed and closed fractures. The developed ANN model (Zazoun, 2013) can predict fracture density by using conventional well log data. The application of artificial intelligence to fractured reservoir indicates that a neural network can be successfully used to predict the fracture density in boreholes using conventional well log data. The results of this study show that conventional well log data such as gamma ray (GR), sonic interval transit time, caliper, neutron porosity and bulk density logs, and the core depth data are suitable inputs to build an ANN modeling.

However, at present, there is a growing need to develop new approaches for well facies classification. More and more attention is being drawn to deep neural networks. The presence of a large number of layers allows the neural network to construct a conception of the research object from simple features gradually moving to more complex ones.

In this paper, an architecture based on a one-dimensional convolutional neural network (1D-CNN) is proposed for facies classification. The photoelectric effect (PE), GR, resistivity logging (RL), neutron-density porosity difference (DPHI), average neutron density porosity (PHIA) and geologic constraining variables are considered as input data of the model.

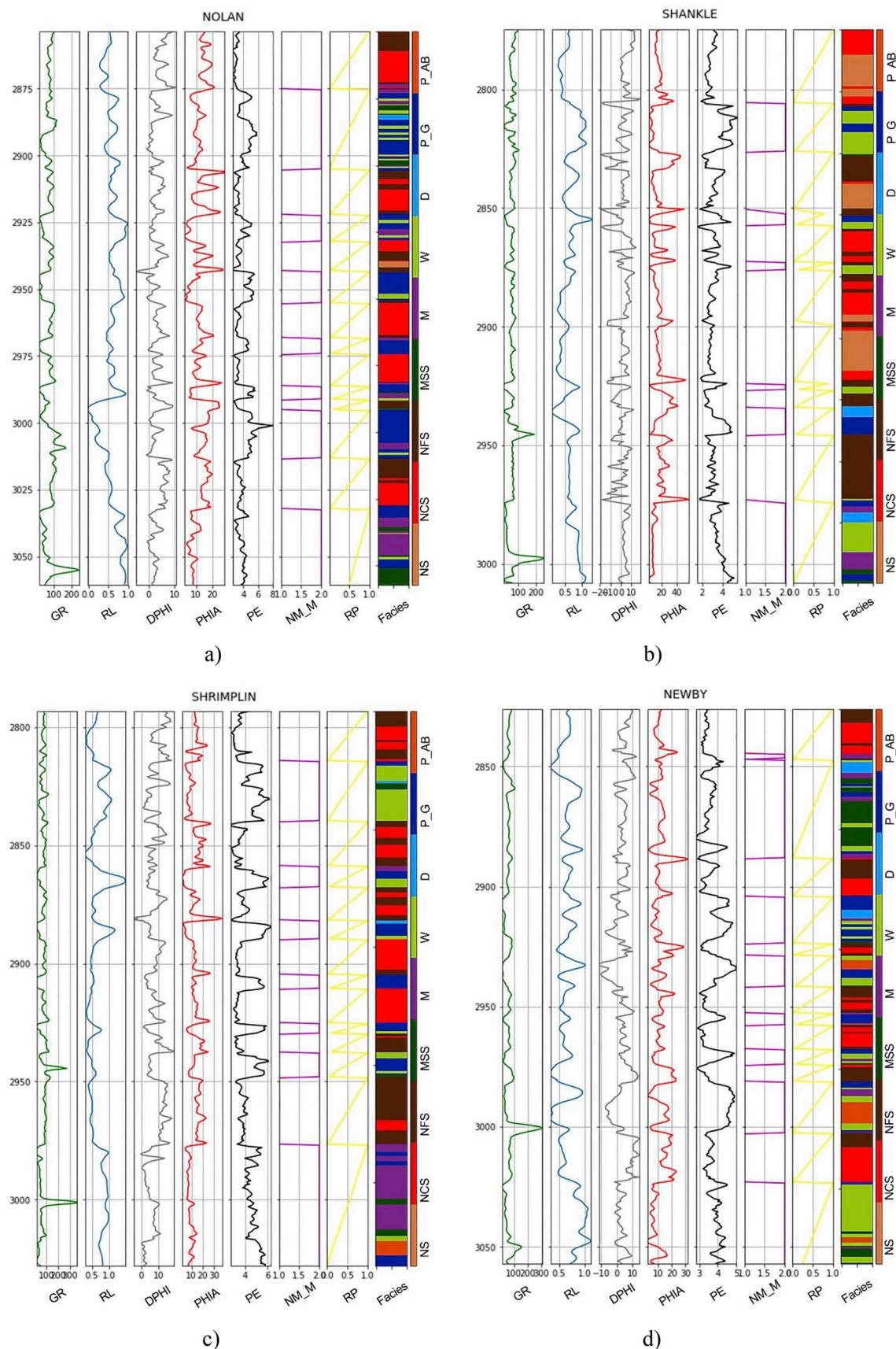


Fig. 3. Facies and seven log curves of ten wells.

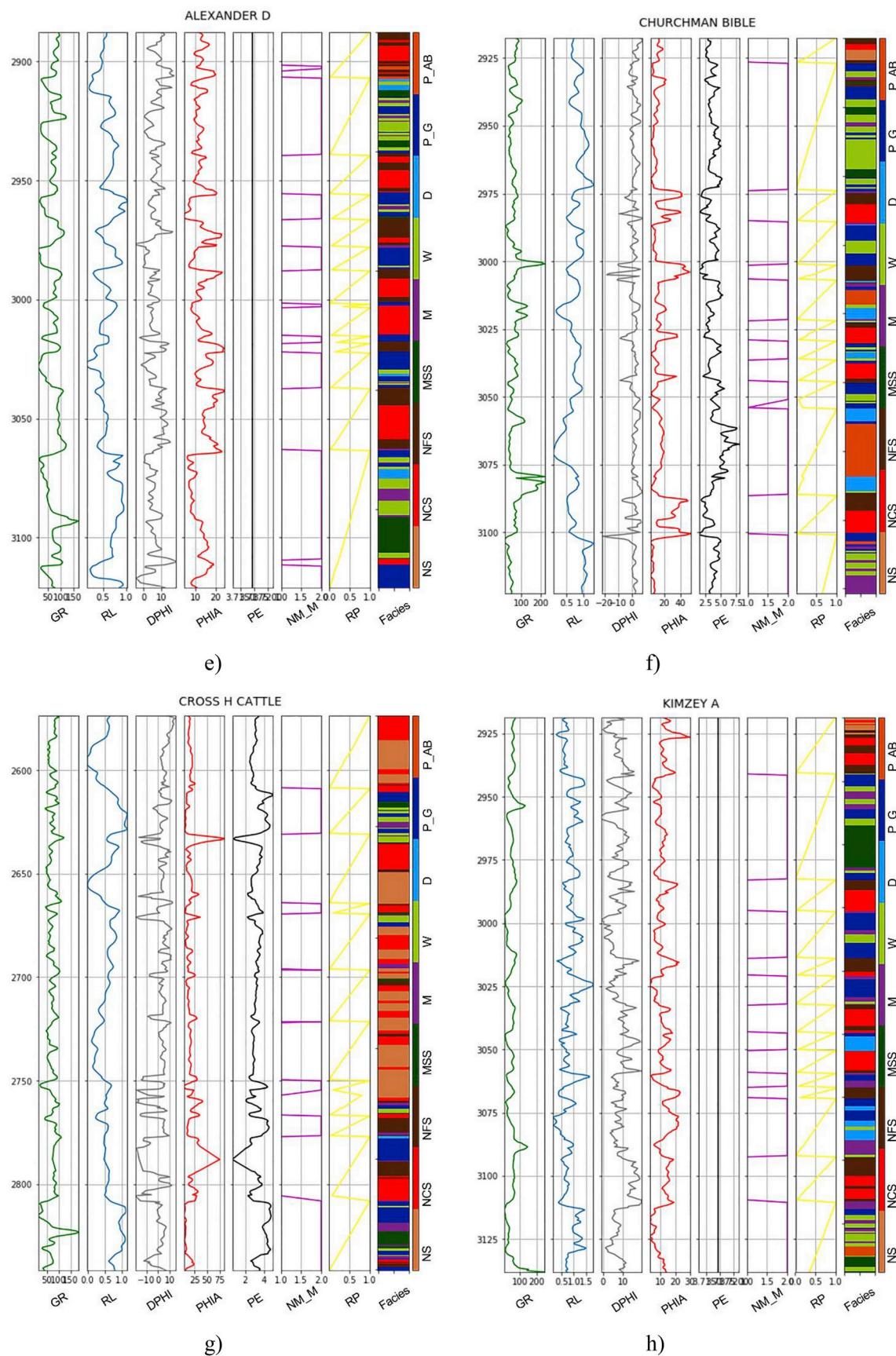


Fig. 3. (continued)

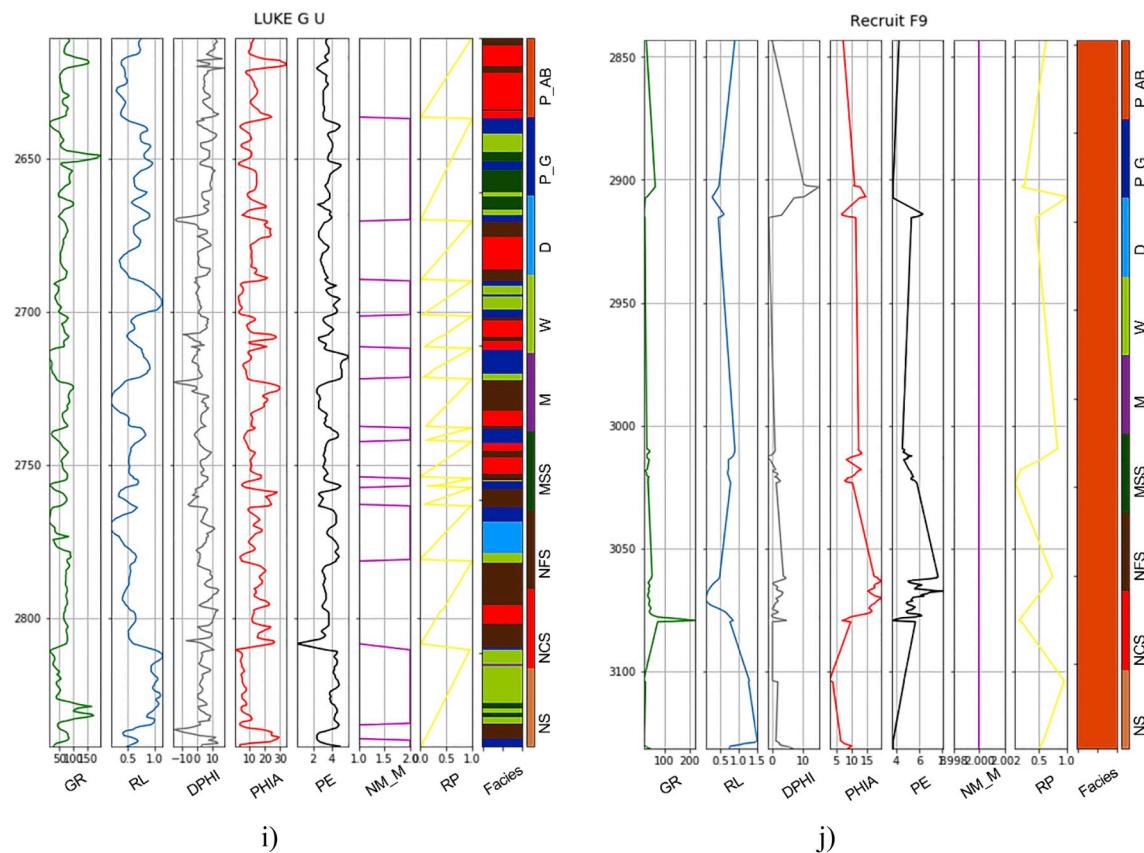


Fig. 3. (continued)

**Table 1**

The color scheme of nine facies.

Facies type	Color	Facies type	Color	Facies type	Color
Nonmarine sandstone	#C8A26E	Marine siltstone and shale	#3A3A5A	Dolomite	#0066CC
Nonmarine coarse siltstone	#E63333	Mudstone (limestone)	#8050A0	Packstone-grainstone (limestone)	#5A5A80
Nonmarine fine siltstone	#664D4D	Wackestone (limestone)	#80B033	Phylloid-algal bafflestone (limestone)	#C88033

The rest of the paper is organized as follows. Section 2 describes the area of the study. Section 3 presents the architecture of the proposed 1D-CNN model. Section 4 evaluates the proposed model using various optimization algorithms to illustrate the benefits of the proposed implementation according to computational efficiency and classification accuracy. Then a discussion of the results is given, followed by conclusions in Section 5.

## 2. Area of study

The research area is a hydrocarbon field located in the southwest of Kansas known as the Hugoton basin, a northern shelf expansion greater and deeper part located under Oklahoma and Texas. For millions of years, thousands of feet of carbonate (limestone and dolomite) deposits and shale have been accumulated.

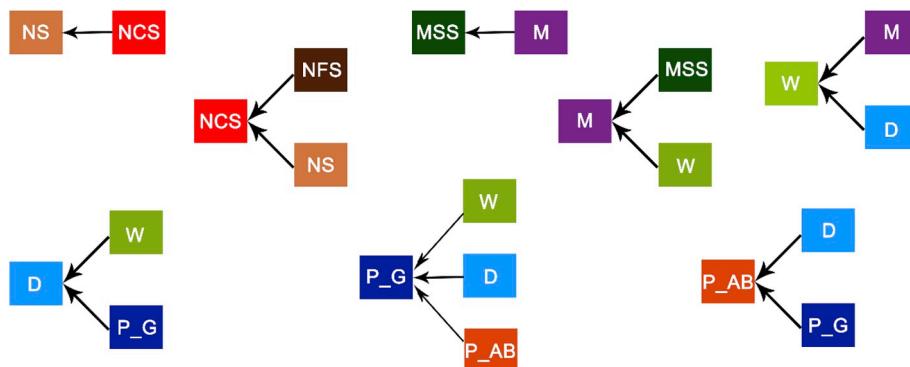
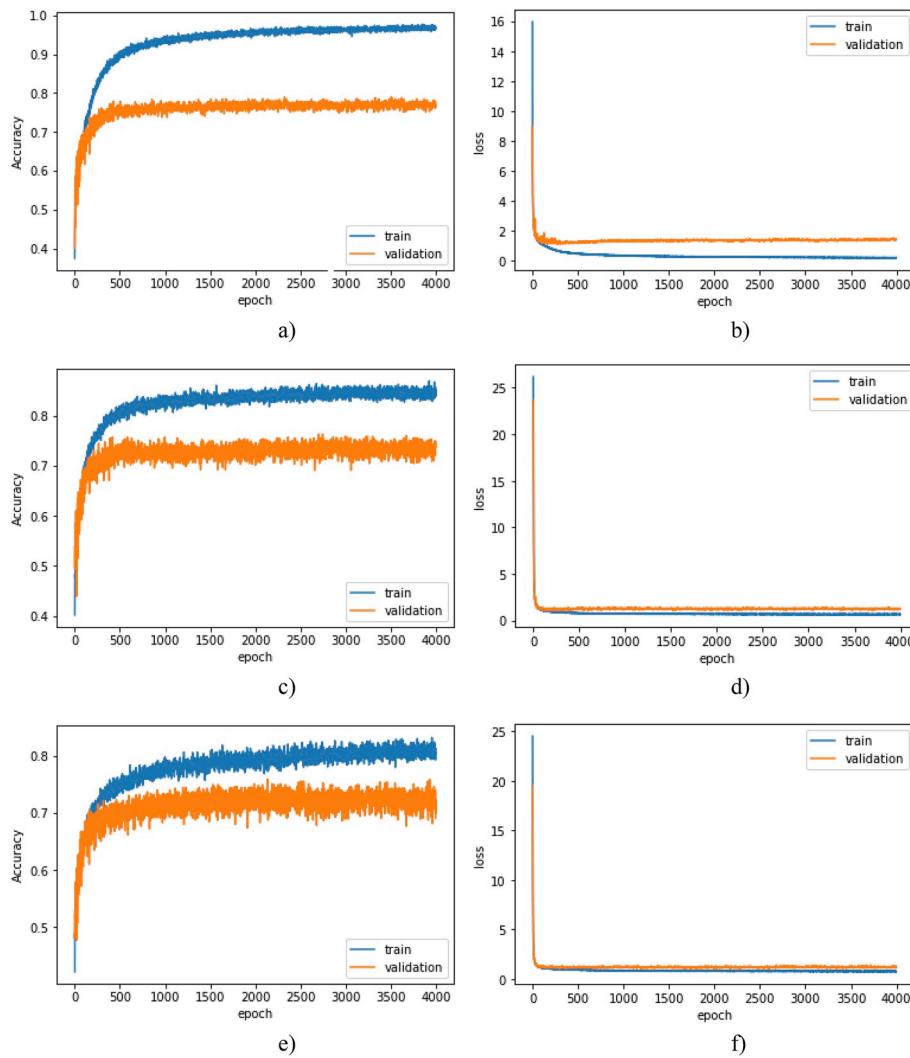


Fig. 4. The description of adjacent facies in wells.

**Table 2**

Evaluation of the proposed approach with RNN and LSTM models.

	Training loss	Validation loss	Training accuracy (%)	Validation accuracy (%)	Adjacent facies classification accuracy for validation dataset (%)
1D-CNN (Adagrad)	0.2173	1.4345	96.44	76.97	93.20
1D-CNN (Adadelta)	0.6331	1.1844	85.65	73.88	93.20
1D-CNN (Adamax)	0.7495	1.1945	81.20	69.09	92.12
RNN	1.0246	1.0995	59.28	57.38	90.36
LSTM	0.7631	1.0289	70.43	62.95	90.48

**Fig. 5.** Facies classification loss and accuracy curves for the training and validation datasets for 1D-CNN(Adagrad) (a,b), 1D-CNN(Adadelta)(c,d) and 1D-CNN(Adamax)(e,f).

The Hugoton gas field makes a significant contribution to the economy of Kansas. The fields constitute the largest gas producing zone in North America, acquiring 963 billion m<sup>3</sup> of gas from 412,000 wells.

Hugoton has the deepest wells in Kansas State reaching up to 5000 feet. Most of the gas is produced by the Chase and Council Grove Groups. The rocks were precipitated during the Permian period (Dubois et al., 2003). Council Grove Group production is limited to the Panoma field, which is under and geographically covered by the Hugoton field (Dubois et al., 2006). The study area includes the location of approximately 515 wells with measured and derived predictor variables (Dubois et al., 2007). The work of this particular field is studied in this paper.

The lithofacies identified in this study are based on a visual examination of a 4149-foot core of nine wells. Physical and chemical

properties of rocks were determined using logging tools. Well logs include five wireline log curves (GR, resistivity logging (RL), photoelectric effect (PE), neutron-density porosity difference (DPHI) and average neutron density porosity (PHIA)) and two geologic constraining variables (nonmarine-marine indicator (NM\_M) and relative position (RP)). Digital measurements are recorded at half-foot (0.15 m) intervals (Dubois et al., 2007).

The analyzed sequence represents by nonmarine sandstone, coarse and fine siltstone, marine siltstone and shale, mudstone, wackestone, dolomite, packstone-grainstone, and phylloid-algal bafflestone.

### 3. Methodology

The lithofacies from the core, associated well log variables, served

**Table 3**

Comparison of facies classification results using the proposed approach with RNN, LSTM, SVM, and KNN.

Method	Evaluation metrics	Facies									
		NS	NCS	NFS	MSS	M	W	D	P_G	P_AB	Total
RNN	Accuracy (%)	75.86	65.00	62.05	58.70	18.18	42.50	46.15	49.72	80.00	56.39
	F-measure (%)	53.01	67.29	64.78	55.10	10.53	43.97	43.64	57.69	52.46	55.22
LSTM	Accuracy (%)	81.25	69.36	68.92	50.85	44.00	50.39	72.73	58.13	83.87	63.13
	F-measure (%)	60.47	74.09	68.00	54.05	27.85	53.94	40.00	63.92	72.22	62.22
SVM	Accuracy (%)	78.00	75.23	76.06	73.47	58.82	62.07	83.33	76.34	<b>88.89</b>	73.73
	F-measure (%)	75.00	78.22	73.47	71.29	57.14	63.16	75.47	76.34	<b>93.02</b>	73.64
KNN	Accuracy (%)	73.08	74.31	74.66	64.62	56.25	58.87	83.33	77.27	76.74	71.33
	F-measure (%)	71.70	76.60	73.15	71.19	52.94	61.86	75.47	70.83	78.57	71.32
1D-CNN (Adagrad)	Accuracy (%)	<b>84.09</b>	<b>77.63</b>	<b>76.82</b>	<b>74.55</b>	<b>70.69</b>	<b>67.89</b>	<b>86.67</b>	<b>78.15</b>	<b>88.89</b>	<b>76.87</b>
	F-measure (%)	75.51	<b>80.38</b>	<b>76.57</b>	71.93	<b>70.69</b>	<b>67.58</b>	<b>80.00</b>	<b>78.81</b>	<b>93.02</b>	<b>76.78</b>
1D-CNN (Adadelta)	Accuracy (%)	79.59	75.77	76.47	<b>74.55</b>	61.36	62.07	84.62	77.50	85.97	74.58
	F-measure (%)	75.00	78.54	74.02	<b>77.36</b>	55.10	64.87	<b>80.00</b>	75.30	89.91	74.44
1D-CNN (Adamax)	Accuracy (%)	78.57	75.23	76.43	73.68	58.62	61.39	84.62	76.61	81.25	73.37
	F-measure (%)	<b>77.19</b>	78.10	73.29	75.00	59.13	58.77	77.19	76.92	88.64	73.20

The best results according to the accuracy and f-measure metrics are marked in bold.

**Table 4**

Comparison of adjacent facies classification results using the proposed approach with RNN, LSTM, SVM, and KNN.

Method	Evaluation metrics	Facies									
		NS	NCS	NFS	MSS	M	W	D	P_G	P_AB	Total
RNN	Accuracy (%)	98.08	99.51	94.94	64.58	83.72	83.62	89.66	84.29	94.87	90.36
	F-measure (%)	96.23	99.51	96.77	62.00	74.23	85.09	89.66	87.08	92.50	90.20
LSTM	Accuracy (%)	92.73	99.51	95.48	56.90	86.67	86.73	<b>100.00</b>	86.33	94.87	90.48
	F-measure (%)	93.58	99.51	96.42	60.00	78.79	87.11	84.00	88.89	92.50	90.50
SVM	Accuracy (%)	94.44	98.54	95.48	78.72	89.29	91.96	<b>100.00</b>	91.67	93.18	93.73
	F-measure (%)	94.44	98.54	96.42	74.75	90.91	91.96	92.59	92.02	96.47	93.67
KNN	Accuracy (%)	92.73	99.03	96.77	70.49	81.36	88.89	<b>100.00</b>	95.00	83.33	92.17
	F-measure (%)	93.58	99.27	97.72	76.11	84.96	87.27	90.57	90.84	84.34	92.24
1D-CNN (Adagrad)	Accuracy (%)	<b>98.18</b>	<b>100.00</b>	<b>97.42</b>	79.25	<b>90.48</b>	92.73	<b>100.00</b>	<b>95.90</b>	<b>95.56</b>	<b>95.54</b>
	F-measure (%)	<b>99.08</b>	99.75	<b>98.37</b>	80.00	<b>93.44</b>	<b>92.73</b>	<b>98.04</b>	92.86	96.63	<b>95.54</b>
1D-CNN (Adadelta)	Accuracy (%)	96.61	99.51	97.26	79.25	89.29	<b>93.52</b>	<b>100.00</b>	95.35	93.75	95.18
	F-measure (%)	97.44	<b>99.76</b>	97.26	<b>84.85</b>	90.09	91.82	89.29	<b>94.25</b>	<b>96.77</b>	95.19
1D-CNN (Adamax)	Accuracy (%)	95.00	99.51	96.67	<b>80.36</b>	87.50	92.45	<b>100.00</b>	94.96	93.48	94.58
	F-measure (%)	96.61	99.51	97.64	81.82	89.60	90.74	92.31	93.00	96.63	94.57

The best results according to the accuracy and f-measure metrics are marked in bold.

as the input data for the training of a deep convolutional neural network model. The model is used to evaluate lithofacies in wells without core data, but having the appropriate logging curves. Testing was carried out in parallel with the model construction and served to verify the properties and the working process of the model.

### 3.1. The workflow of the proposed approach

Fig. 1 shows the general workflow of a deep convolutional network classifier. The development of the classifier consists of the following steps:

- 1) Obtaining well logs and building training set and testing set to verify the performance of the model.
- 2) Data preprocessing.
- 3) Construction of a deep convolutional neural network as a classifier of well facies. Determine the number of neurons for each hidden layer.
- 4) Training of the classifier on the well log data, taking into account the core data to obtain the optimal solution.
- 5) Evaluation of the obtained classifier, its application to an unknown log dataset.

### 3.2. The architecture of the proposed model

The architecture of the proposed 1D-CNN model consists of an input

layer, four convolutional layers with ReLU (a rectified linear unit) as a nonlinear activation function, two maxpooling layers and three fully-connected layers. The last layer is the output layer, which assigns a label to the input data. Below is a detailed explanation of the proposed architecture (Fig. 2).

The input vector is given to the input layer and is then propagated through several successive convolutional and pooling layers to extract features. Pooling layers are used for a subsample to reduce the network dimension, which can help reduce computation and control overfitting. Each convolution layer has several one-dimensional convolutional filters (kernels). The size of the kernel is a hyperparameter and depends on the data.

Let us denote the input data vector as  $X = (x_1, x_2, \dots, x_N) \in R^n$ , where  $n$  is the length of the input vector. In the first convolutional layer, a set of  $d$  filters  $\Sigma = (\sigma_1, \sigma_2, \dots, \sigma_d)$  is given to the input vector through the convolution operation ( $\cdot$ ) to obtain a feature map:

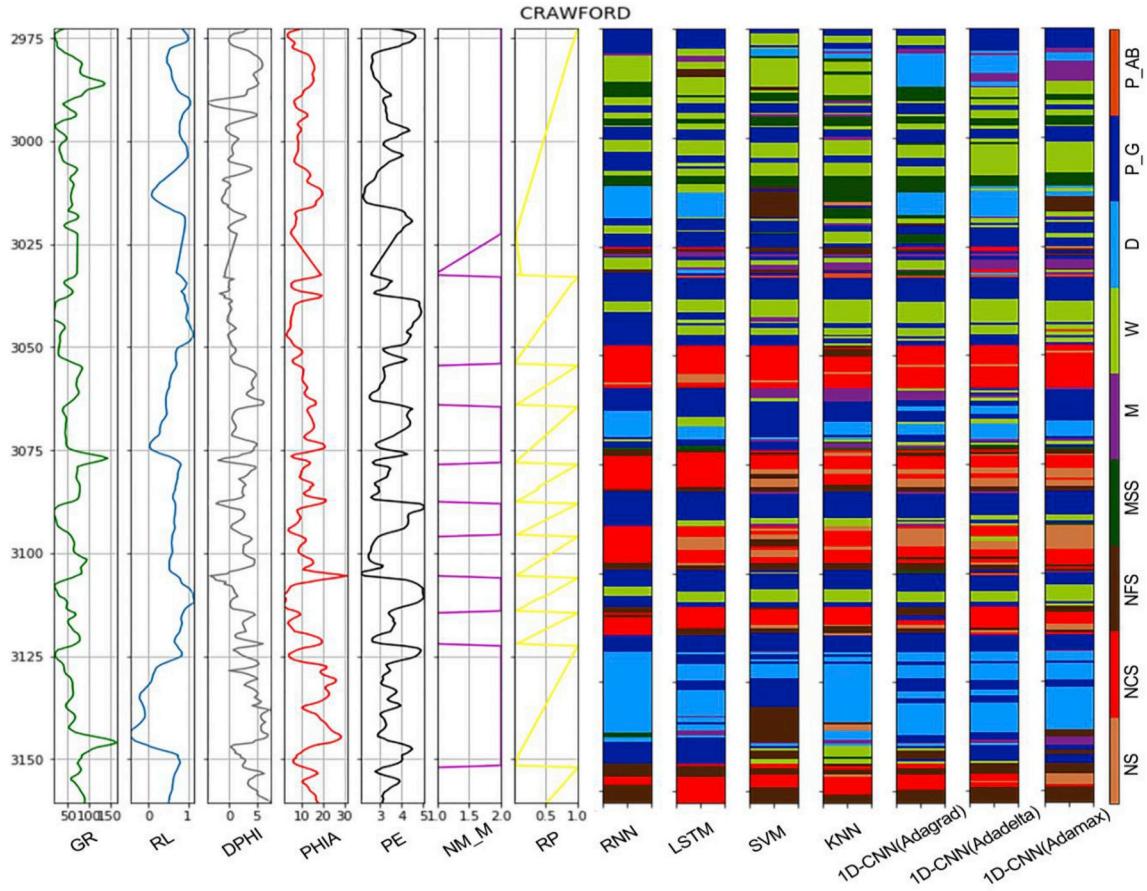
$$F = (f_1, f_2, \dots, f_N) = g(X \cdot \Sigma) \quad (1)$$

where  $g$  is a nonlinear activation function, such as hyperbolic tangent or ReLU. ReLU is defined as

$$g(x) = \max(0, x) \quad (2)$$

It became the most used activation function in CNN.

After applying the ReLU function, the feature maps are sent to the second convolution layer with filters of the same size as in the first convolutional layer. The generated feature maps are then sent to the



**Fig. 6.** Evaluation results of the CRAWFORD well.

first maxpooling layer. The maxpooling layers reduce the samples of output features of the previous layer for better control of high-dimensional data.

The most common form is a pooling layer with filters of length 2. After applying this layer, the depth (size) remains unchanged, but the length is reduced by half.

The resulting output volume is sent to the third convolution layer in the proposed architecture. This layer is intended to further refine the feature maps by processing each element one by one. Next comes the fourth convolutional layer. Again, after applying the ReLU function, the feature maps generated on the previous layer are sent to the second maxpooling layer, with the  $2 \times 1$  kernel.

The extracted high-level features will then be flattened to a vector of fixed size to pass through the fully connected layers. 2500 and 1500 indicate the number of units in the first and second fully connected layers, respectively. These two layers calculate their output as

$$y_k = f(w_k y_{k-1} + b_k) \quad (3)$$

where  $w_k$  are weight matrices,  $b_k$  are bias vectors,  $y_{k-1}$  is the output of the previous layer, and the activation function  $f(\cdot)$  is ReLU,  $k = 1, 3$ . Further, the output of the second fully connected layer is sent to the third fully connected layer, which uses the softmax activation function to calculate the predictive probabilities for all classes.

This can be achieved by:

$$p(\hat{y} = j|x) = \frac{\exp(w_j^T x + b_j)}{\sum_{i=1}^K \exp(w_i^T x + b_i)} \quad (4)$$

where  $w_j$  and  $b_j$  are parameters of the softmax function for the  $j^{\text{th}}$  class,  $\hat{y}$  is the predicted class, and  $K$  is the number of classes (in our case  $K = 9$ ).

An additional and customized screening mechanism was applied to

the second and fourth convolutional layers, as well as to the first fully connected layer, to avoid the overfitting problems. The dropout method discards some randomly selected hidden neurons, and they are not used in the back-propagation stage (Hinton et al., 2012).

Batch normalization layers usually follow convolutional layers to normalize output. It allows to increase the training speed significantly and makes the network less sensitive to initialization (Ioffe and Szegedy, 2015). We also applied batch normalization after the first fully connected layer.

Then the network is trained. The neural network training is aimed at finding the best parameters (network weights) to minimize the loss function, which, in the classification problem, measures the compatibility between the prediction (for example, class scores in the classification) and the true label.

A cross-entropy loss is considered to determine the loss of the proposed 1D-CNN model. It is defined as follows (Ceci et al., 2017):

$$L = -\frac{1}{N} \sum_{i=1}^N y_i \cdot \log(p(\hat{y}_i)) \quad (5)$$

where  $y_i$  is the true class of the  $i^{\text{th}}$  sample,  $N$  is the number of training samples.

The entire consecutive network (all weights and biases) can be optimized by an optimization algorithm using cross-entropy as training metric. Adagrad (Duchi et al., 2011), Adadelta (Zeiler, 2012), and Adamax (Kingma and Ba, 2015) are used in this work.

#### 4. Experimental results and discussion

The 1D-CNN model is implemented in Python 2.7.13 using various libraries, including Tensorflow and Keras. All experiments were conducted on Intel Xeon (R), CPU X5670 @ 2.93 GHz  $\times 4$  with 10 GB of

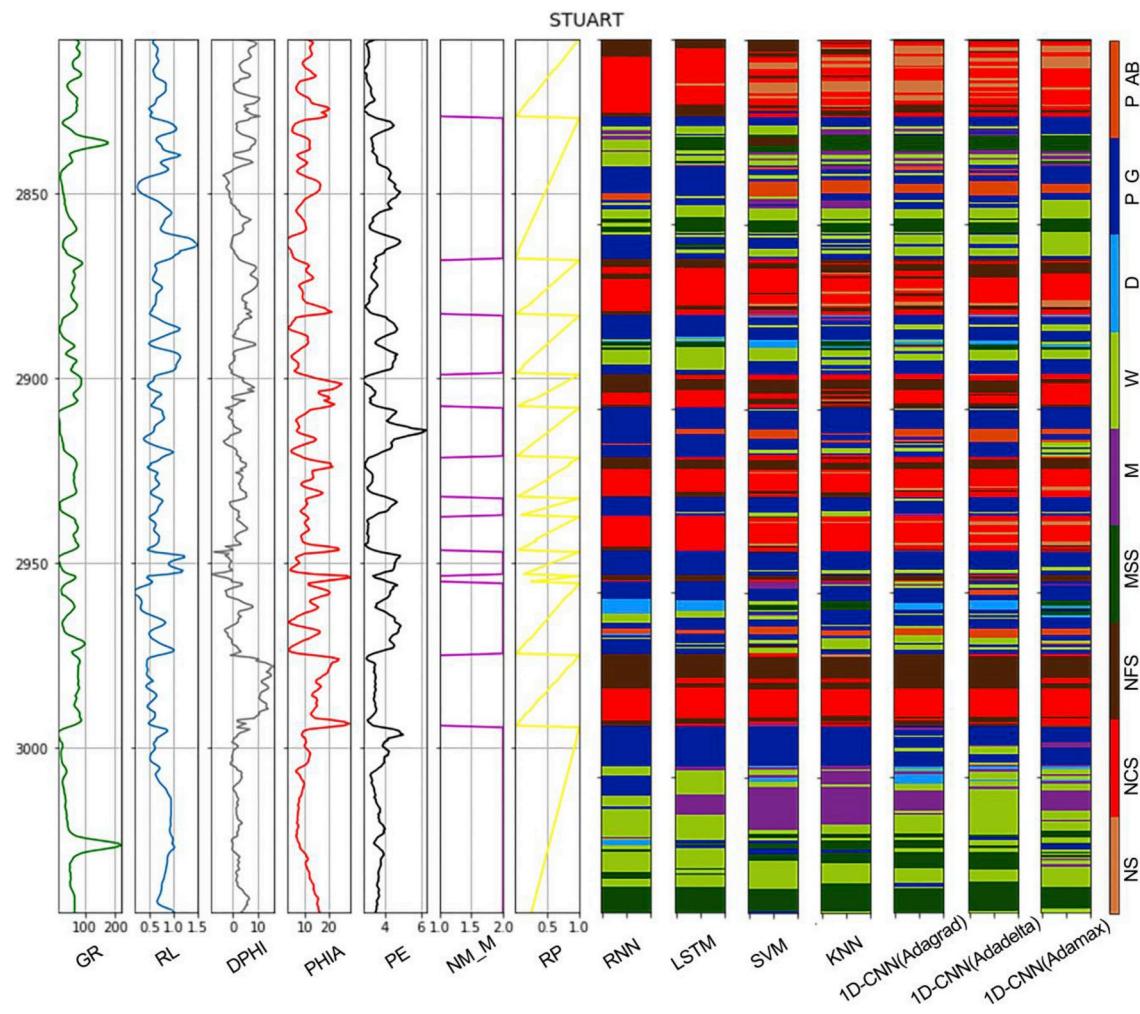


Fig. 7. Evaluation results of the STUART well.

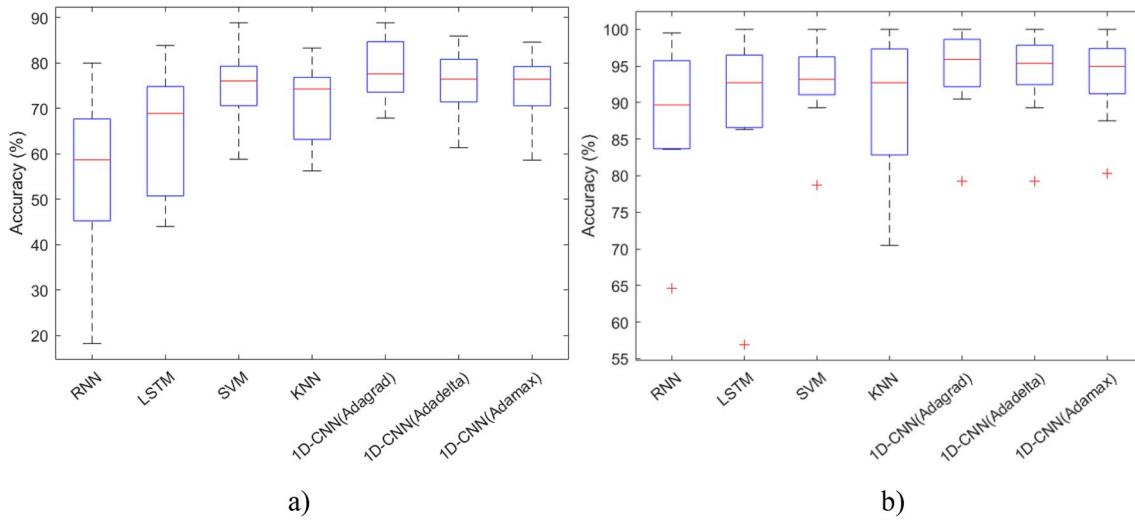


Fig. 8. Boxplot diagram of the performance evaluation of classification models.

RAM machine.

We use well log data to train the proposed 1D-CNN model to classify facies. In the considered dataset, there are data from nine wells containing 4149 samples. A pseudo-well “Recruit F9” was introduced to improve the ninth facies (phyllloid-algal bafflestone, P\_AB) classification accuracy (Dubois et al., 2007). It was necessary since other wells

contain only a few P\_AB facies, which makes it difficult to classify them. Thus, the dataset uses log data from 10 wells (SHRIMPLIN, ALEXANDER D, SHANKLE, LUKE G U, KIMZEY A, CROSS H CATTLE, NOLAN, Recruit F9, NEWBY, and CHURCHMAN BIBLE) that were labeled with a facies type based on the core data (Fig. 3).

The feature vector is constructed using the following physical

**Table 5**

Parameters of the boxplot.

	RNN	LSTM	SVM	KNN	1D-CNN (Adagrad)	1D-CNN (Adadelta)	1D-CNN (Adamax)
Minimum	18.18	44.00	58.82	56.25	<b>67.89</b>	61.36	58.62
$Q_1$	18.18	44.00	58.82	56.25	<b>67.89</b>	61.36	58.62
Median	58.70	68.92	76.06	74.31	<b>77.63</b>	76.47	76.43
Mean	55.35	64.39	74.69	71.01	<b>78.38</b>	75.32	74.04
$Q_3$	67.72	74.86	79.33	76.87	<b>84.74</b>	80.85	79.24
Maximum	80.00	83.87	<b>88.89</b>	83.33	<b>88.89</b>	85.97	84.62

	RNN	LSTM	SVM	KNN	1D-CNN (Adagrad)	1D-CNN (Adadelta)	1D-CNN (Adamax)
Minimum	64.58	56.90	78.72	70.49	79.25	79.25	<b>80.36</b>
$Q_1$	64.58	56.90	78.72	70.49	<b>79.25</b>	79.25	<b>80.36</b>
Median	89.66	92.73	93.18	92.73	<b>95.90</b>	95.35	94.96
Mean	88.14	88.80	92.59	89.73	<b>94.39</b>	93.84	93.33
$Q_3$	95.73	96.49	96.25	97.34	<b>98.64</b>	97.82	97.38
Maximum	99.51	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>	<b>100</b>

(a)

(b)

The best results according to the accuracy and f-measure metrics are marked in bold.

properties: GR, RL, PE, DPHI, PHIA and two geologic constraining variables (NM\_M and RP). Measured properties and their transformations are potential elements of feature vectors.

The dataset contains the following facies types: nonmarine sandstone (NS), nonmarine coarse siltstone (NCS), nonmarine fine siltstone (NFS), marine siltstone and shale (MSS), mudstone (limestone) (M), wackestone (limestone) (W), dolomite (D), packstone-grainstone (limestone) (P\_G) and P\_AB (limestone). Two facies are of continental origin, namely NCS and NFS, and six of them are of marine origin (MSS, M, W, D, P\_G, and P\_AB). In the non-marine area, NCS typically has a higher permeability than NFS for a given porosity. The input color map for the nine rocks is presented in [Table 1](#) to establish a lithological description of the interpreted well.

The standardization procedure was applied to the data so that it had a zero mean and unit variance. PE feature is not available for all wells. It is available for only 3232 samples out of 4149 possible. We replaced missing values with mean values of the PE column.

The proximity of facies with each other leads to their mixing. The facies and their approximate neighbors are indicated and marked by abbreviations and consistent colors in [Fig. 4](#). Experiments are also conducted on a data from unknown wells that do not have facies labels.

The first step in the experiments was to change the configuration parameters of the convolutional neural network to obtain the best possible classification accuracy in the dataset by cross-validation. The input data is divided into training and validation sets (20% of the dataset) to conduct the experiments. Batch size was determined as 10, and the loss function was chosen as the categorical cross-entropy. Each experimental result was obtained over 4000 epochs to provide consistent comparisons.

The proposed model is trained on various optimization methods, such as Adagrad, Adadelta, and Adamax and is denoted as 1D-CNN (Adagrad), 1D-CNN(Adadelta) and 1D-CNN(Adamax), respectively. It is compared with such neural networks as the recurrent neural network (RNN) ([Jordan, 1986](#)) and long-short term memory (LSTM) ([Hochreiter and Schmidhuber, 1997](#)). The results are compared and shown in [Table 2](#).

The accuracy (which reflects the prediction ability of the model) and the loss function's tendencies when applying the three optimization methods to the proposed model are shown in [Fig. 5](#).

The figure shows the cross-entropy error of the proposed CNN model as a function of the iteration number when different optimizers are used. It decreases with the increase of iteration number when all three optimizers are used, which shows quite a high accuracy of the

model.

In comparison with the other two models, 1D-CNN(Adagrad) is the most stable (with the increase of iteration number). Thus, a model trained with Adamax fails in this experiment. Its training and validation accuracies are 81.20% and 69.09%, respectively. However, the gap for 1D-CNN(Adagrad) is 19.47%, in contrast to the other two models (11.77% for 1D-CNN(Adadelta) and 12.11% for 1D-CNN(Adamax)). The use of Adagrad results in the highest validation and training accuracies.

The accuracy of both training and validation datasets using 1D-CNN (Adamax) is the lowest but showed a higher result than RNN and LSTM. In general, the model, trained with the Adagrad optimizer, surpasses the models trained by Adamax and Adadelta.

A comparative analysis of the proposed model using Adagrad, Adadelta and Adamax optimizers with RNN, LSTM, SVM ([Cortes and Vapnik, 1995](#)) ([Hall, 2016](#)) and k-nearest neighbors (KNN) ([Cover and Hart, 1968](#)) based on accuracy and F-measure metrics is shown in [Table 3](#). The best results are marked in bold for a more detailed evaluation of the classification characteristics of the proposed model at the end of the 4000<sup>th</sup> epoch.

According to [Table 3](#), the best results were obtained for the classification of NS, NCS, D, P\_G, and P\_AB (84.09%, 77.63%, 86.67%, 78.15%, and 88.89%, respectively) for 1D-CNN (Adagrad). And the accuracy of M and W rocks classification led to the lowest results (70.69% and 67.89%, respectively).

After considering RNN, LSTM, SVM, and KNN models and comparing them with each other, it was obtained that the best accuracy is achieved for NCS, NFS, MSS, M and W facies (75.23%, 76.06%, 73.47%, 58.82 and 62.07%, respectively) using SVM model. But for NS and P\_AB facies it yielded to LSTM, and for P\_G facies - KNN.

At the same time, the classification accuracy of the MSS facies for 1D-CNN(Adagrad) and 1D-CNN(Adadelta) coincided (74.55%), and the F-measure showed a higher result for the second model. According to the F-measure values for the D facies, both models showed the same result.

[Table 4](#) gives the results of the adjacent facies classification. Applying Adagrad to the proposed 1D-CNN model for the NS, NCS, NFS, M, D, P\_G and P\_AB adjacent facies resulted in 98.18%, 100%, 97.42%, 90.48%, 100%, 95.90%, and 95.56%, respectively, according to the accuracy. According to the F-measure, the proposed model showed the best result. The use of the Adamax optimizer in comparison with Adagrad and Adadelta increased the classification accuracy of the MSS facies (80.36%).

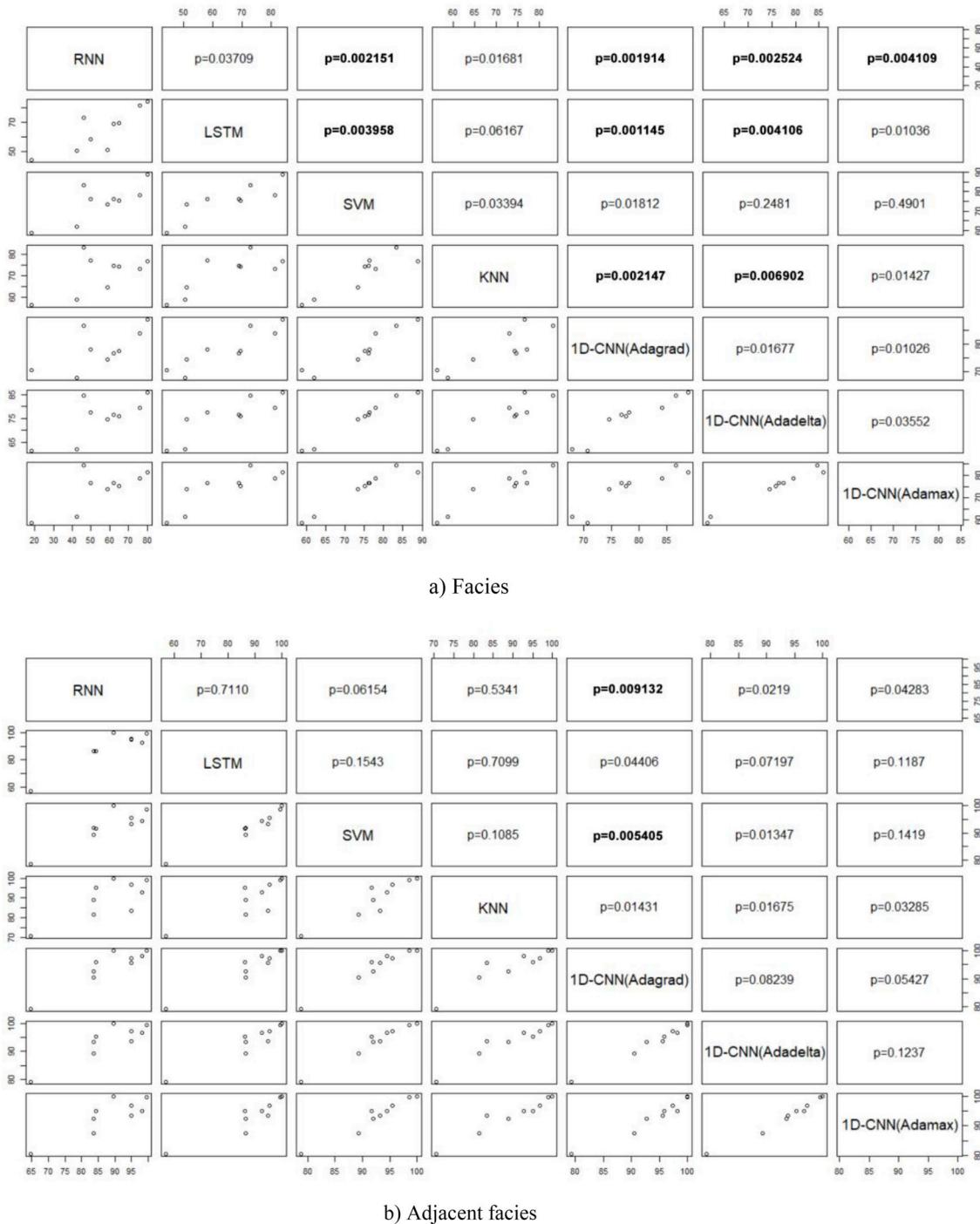


Fig. 9. Pairwise comparison between models.

According to the table, rather good accuracy results for the classification of adjacent facies showed RNN for NS, NCS, D, and P\_AB facies; LSTM for NCS and P\_AB; SVM for MSS, M, and W, and KNN for NFS facies and P\_G. The D facies were accurately identified and showed a 100% result for the proposed model, and the LSTM, SVM and KNN models. The F-measure, as a common measure of the relevance of the classifier, correctly classifies NS with a classification accuracy of 99.08%, NFS - 98.37% and D - 98.04%.

In summary, the accuracy of the validation data for the optimizers Adamax and Adadelta are close and inferior to Adagrad. Consequently, Adagrad is superior to Adadelta and Adamax in this experiment.

A set of two wells (830 samples, where 474 refer to STUART, and

356 to CRAWFORD) which do not have labels is used to test the trained model. The obtained results can be estimated from the well log data.

Logging curves have a stratigraphic meaning. It is possible to determine sections of well logs with physical significance. Therefore, zoning according to well logs is a very important step in the pre-processing of well data.

Comparison of the results obtained in the peaks of the log curves makes it possible to evaluate the correctness of the proposed model in the corresponding zone of the well. By the nature of the deflections of the DPHI and PHIA curves, the permeability and porosity of formations can be determined. The remoteness of the curves from each other indicates the type of fluid. High resistivity and permeability may indicate

the content of hydrocarbons.

**Fig. 6** shows that the proposed model with three optimizers (Adagrad, Adadelta, and Adamax), together with the considered models accurately determine the location of the NCS, D, P\_G and W rocks for the CRAWFORD well.

As for the STUART well (**Fig. 7**), the results of all models are almost consistent with the detection of P\_G, NCS, MSS, NFS, and W facies.

According to **Fig. 7**, the proposed model with the Adagrad optimizer clearly distinguishes shale from sand, as well as LSTM and SVM. It is confirmed by the reaction of GR (first curve). The curve receives high values for shale and low for sand. For the CRAWFORD well (**Fig. 6**), 1D-CNN(Adagrad) detects the shale along with the RNN model. The high values of the second curve RL indicate the hydrocarbon zone and the low values for the water-saturated zone.

For the CRAWFORD well, the DPHI and PHIA curves in the section from 3000 to 3025 feet are placed sufficiently close to each other, suggesting the presence of limestone. The PE curve in this section approaches 5 b/electron. It is confirmed by the proposed model (1D-CNN (Adagrad) and 1D-CNN(Adadelta)), unlike SVM, which assumes the presence of NFS.

The DPHI and PHIA curves in conjunction with PE sufficiently demonstrate the presence of limestone for the STUART well (**Fig. 7**). In this case, the high PE peaks (for example, at a depth of 2900–2950 feet) are the response to iron. High values of GR and average values of PE characterize the areas with MSS for both wells.

Low GR, relatively low DPHI and high PHIA values characterize the presence of D facies, which is confirmed by the results of the proposed model for the Adagrad and Adadelta optimizers. Sections with low PHIA indicate the presence of W facies. The experimental results show that the all considered models classify the marine and non-marine zones quite accurately, which is confirmed by the NM\_M curve.

The boxplot diagrams are given below for a visual comparison of the considered models' performances for facies (**Fig. 8(a)**) and adjacent facies (**Fig. 8(b)**) classification. It can be seen that the maximum accuracy and the highest minimum accuracy are observed for 1D-CNN (Adagrad).

The proposed model also surpassed SVM for more than 50% of the results. The worst performance, according to the boxplot (**Fig. 8(a)**), is observed for RNN. According to **Fig. 8 (b)** 1D-CNN(Adagrad) and 1D-CNN(Adadelta) outperformed SVM by maximum accuracy and 1D-CNN (Adamax) by medium accuracy. According to the analysis mentioned above, it can be concluded that the proposed 1D-CNN(Adagrad) model provides significantly high accuracy for well facies classification.

The visual analysis shows the advantages of applying the proposed approach. **Table 5** lists the parameters used in the boxplot construction (**Fig. 8**), where  $Q_1$  and  $Q_2$  are the first and third quartile, respectively. The best values are marked in bold.

The following shows the pairwise matrices of the seven models when classifying facies (**Fig. 9 (a)**) and adjacent facies (**Fig. 9 (b)**). The lower triangle shows the correlations of the differences of the accuracy for each pair of approaches. The upper triangle represents the corresponding  $p$ -values of the correlated  $t$ -test ([de Assis Boldt et al., 2017](#)). The level of significance is equal to 0.05 ( $\alpha = 0.05$ ). The names of the models are indicated on the main diagonal.

**Fig. 9** shows the differences between the considered models. A statistically significant difference is highlighted in bold. The smallest  $p$ -value (**Fig. 9 (a)**) is equal to 0.001145, and the largest is 0.006902. According to **Fig. 9 (b)** the largest  $p$ -value = 0.009132, and the smallest  $p$ -value = 0.005405 for the adjacent facies classification.

In other words, the proposed 1D-CNN(Adagrad) model is statistically significantly better than the RNN, LSTM and KNN (**Fig. 9 (a)**). And for the adjacent facies classification, 1D-CNN(Adagrad) is statistically significantly better than RNN and SVM.

## 5. Conclusions

Many studies have been focused on facies classification using well logs. At present, this problem is solved using various machine learning methods. An approach based on CNN was proposed in this paper as an effective method for facies classification based on well logging measurements. The proposed model was validated with a real dataset collected and organized in conjunction with the Hugoton and Panoma modeling project ([Dubois et al., 2003](#)) ([Dubois et al., 2006](#)). The model was evaluated using Adagrad, Adadelta and Adamax optimizers. The PE, GR, RL, DPHI, PHIA and geologic constraining variables are considered as features. The results showed that the best network response predictions were obtained for the Adagrad case. Moreover, the 1D-CNN model showed more accurate results compared to SVM, KNN, RNN, and LSTM. Application of the proposed approach to the facies classification showed significant results for marine origin (dolomite, phylloid-algal bafflestone, and packstone-grainstone) and continental origin (coarse siltstone and sandstone) facies. The case study presented and considered in this paper proves the high efficiency of the proposed 1D-CNN model. The proposed 1D-CNN(Adagrad) model showed a statistically significant improvement in the classification of facies and adjacent facies. It is suitable for lithological identification of complex geological structures. The proposed approach can be considered as a preliminary stage of quantitative petrophysical analysis of wells. Summarizing the above, an analysis of the evaluation of the proposed model based on deep learning can be useful for future research and facies identification.

## Acknowledgments

This work was supported by the SOCAR Science Foundation of Azerbaijan.

## References

- Adibifard, M., Tabatabaei-Nejad, S.A.R., Khodapanah, E., 2014. Artificial neural network (ANN) to estimate reservoir parameters in naturally fractured reservoirs using well test data. *J. Petrol. Sci. Eng.* 122, 585–594.
- Ahmadi, M.A., Soleimani, R., Lee, M., Kashwao, T., 2015. Determination of oil well production performance using artificial neural network (ANN) linked to the particle swarm optimization (PSO) tool. *Petroleum* 1 (2), 118–132.
- Auda, G., Kamel, M.S., 1999. Modular neural networks a survey. *Int. J. Neural Syst.* 9 (2), 129–151.
- Ayala, L.F., Ertekin, T., 2005. Analysis of gas-cycling performance in gas/condensate reservoirs using neuro-simulation. In: Paper SPE 95655 Presented at the SPE Annual Technical Conference and Exhibition, Dallas, Texas.
- Ceci, M., Hollmén, J., Todorovski, L., Vens, C., Dzeroski, S., 2017. Machine learning and knowledge discovery in databases - European conference, ECML PKDD 2017, Skopje, Macedonia, September 18–22, 2017. Proc. Part I. Lect. Notes Comput. Sci. 10534 Springer.
- Chehrazi, A., Rezaee, R., 2012. A systematic method for permeability prediction, a Petro-Facies approach. *J. Petrol. Sci. Eng.* 82–83, 1–16.
- Cortes, C., Vapnik, V., 1995. Support-vector networks. *Mach. Learn.* 20, 273. <https://doi.org/10.1007/BF00994018>.
- Cover, T.M., Hart, P.E., 1968. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theor.* 13, 21–27.
- de Assis Boldt, F., Rauber, T.W., Oliveira-Santos, T., Rodrigues, A., Varejão, F.M., Ribeiro, M.P., 2017. Binary feature selection classifier ensemble for fault diagnosis of submersible motor pump. In: IEEE 26th International Symposium on Industrial Electronics (ISIE), Edinburgh, UK.
- Dubois, M.K., Bohling, G.C., Chakrabarti, S., 2007. Comparison of four approaches to a rock facies classification problem. *Comput. Geosci.* 33 (5), 599–617. <https://doi.org/10.1016/j.cageo.2006.08.011>.
- Dubois, M.K., Byrnes, A.P., Bohling, G.C., Doveton, J.H., 2006. Multiscale geologic and petrophysical modeling of the giant Hugoton gas field (Permian), Kansas and Oklahoma. In: Harris, P.M., Weber, L.J. (Eds.), Giant Reservoirs of the World: from Rocks to Reservoir Characterization and Modeling. American Association of Petroleum Geologists Memoir, pp. 307–353.
- Dubois, M.K., Byrnes, A.P., Bohling, G.C., Seals, S.C., Doveton, J.H., 2003. Statistically-based lithofacies predictions for 3-D reservoir modeling: examples from the Panoma (Council Grove) field, Hugoton embayment, southwest Kansas. In: American Association of Petroleum Geologists 2003 Annual Convention, Salt Lake City, Utah.
- Duchi, J., Hazan, E., Singer, Y., 2011. Adaptive subgradient methods for online learning and stochastic optimization. *J. Mach. Learn. Res.* 12, 2121–2159.

- Elkatatny, S.M., Mahmoud, M., 2018. Development of new correlations for the oil formation volume factor in oil reservoirs using artificial intelligent white box technique. *Petroleum* 4 (2), 178–186. <https://doi.org/10.1016/j.petm.2017.09.009>.
- Gao, P., Jiang, C., Huang, Q., Cai, H., Luo, Z., Liu, M., 2016. Fluvial facies reservoir productivity prediction method based on principal component analysis and artificial neural network. *Petroleum* 2 (1), 49–53.
- Hall, B., 2016. Facies classification using machine learning. *Lead. Edge* 35, 906–909.
- Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.R., 2012. Improving Neural Networks by Preventing Co-adaptation of Feature Detectors. *CoRR* [arXiv:1207.0580].
- Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. *Neural Comput.* 9 (8), 1735–1780.
- Ioffe, S., Szegedy, C., 2015. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: 32nd International Conference on Machine Learning, Lille, France.
- Jordan, M.I., 1986. Serial Order: a Parallel Distributed Processing Approach. Technical Report 8604. Institute for Cognitive Science, University of California, San Diego.
- Kingma, D.P., Ba, J., 2015. Adam: a Method for Stochastic Optimization. [arXiv:1412.6526](https://arxiv.org/abs/1412.6526) [cs.LG].
- Nakutnyy, P., Asghari, K., Torn, A., 2008. Analysis of water flooding through application of neural networks. In: Canadian International Petroleum Conference, Calgary, Alberta.
- Ohl, D., Rae, A., 2014. Rock formation characterization for carbon dioxide geosequestration: 3D seismic amplitude and coherency anomalies, and seismic petrophysical facies classification, Wellington and Anson-Bates Fields, Kansas, USA. *J. Appl. Geophys.* 103, 221–231.
- Saljooghi, B.S., Hezarkhani, A., 2014. Comparison of WAVENET and ANN for predicting the porosity obtained from well log data. *J. Petrol. Sci. Eng.* 123, 172–182.
- Serra, O., 1984. Fundamentals of Well-log Interpretation: the Interpretation of Logging Data. Elsevier, Amsterdam.
- Sfidari, E., Kadkhodaie-Ikhchi, A., Rahimpour-Bonab, H., Soltani, B., 2014. A hybrid approach for litho-facies characterization in the framework of sequence stratigraphy: a case study from the South Pars gas field, the Persian Gulf basin. *J. Petrol. Sci. Eng.* 121, 87–102.
- Wang, X., Yang, S., Zhao, Y., Wang, Y., 2018. Improved pore structure prediction based on MICP with a data mining and machine learning system approach in Mesozoic strata of Gaoqing field, Jiyang depression. *J. Petrol. Sci. Eng.* 171, 362–393.
- Zazoun, R.S., 2013. Fracture density estimation from core and conventional well logs data using artificial neural networks: the Cambro-Ordovician reservoir of Mesdar oil field, Algeria. *J. Afr. Earth Sci.* 83, 55–73.
- Zeiler, M.D., 2012. Adadelta: an Adaptive Learning Rate Method. [arXiv: 1212.5701v1](https://arxiv.org/abs/1212.5701) [cs.LG].