Egyptian Computer Science Journal (ISSN: 1110 – 2586)
Volume 40 – Issue 03, September 2016

A Deep Belief Network Solution for Building a Diminutive Geological Model

[1]Dalia Ahmed Magdi, [2]M. Y. Elgendy
[1]Computer and Information System Department, Sadat Academy for Management Sciences, Cairo, Egypt
[2]Research and Development Department, Apex Company, Cairo, Egypt
daliamagdi@gmail.com, m.elgendy@com-apex.com

Abstract

One of the major current and future challenges facing the oil and gas industry is to maximize the oil recovery factor. One optimization technique is Horizontal wall-bore drilling methodology, "Geo-steering". This type of techniques depends on adjusting the borehole position on-the-fly to reach one or more geological targets; these adjustments are based on the petro-physical information gathered during drilling requires real-time drilling directions decisions. A main step needed for this methodology is building a geological model before drilling for the specified regions based on surrounding pre-drilled wells. In this paper, a proposed solution based on "Deep Belief Network" is presented to construct the geological model and propose a drilling path to follow. In the presented work, a deep architecture "Deep Belief Network" (DBN) with a fine tuning training is used to construct the geological model for a new well before drilling starts. The proposed solution has shown superiority on the traditional Neural Network for determining constructing the model. The proposed system generated a proposed solution achieved more than 92% accuracy based on real training drilling measurements from surrounding 5 wells.

Keywords: *Oil well-bore, Drilling path, Artificial Neural Network (ANN), Deep Learning, Deep Belief Network (DBN).*

1. Introduction

One practical technique to optimize the oil and gas production is to drill non-vertical wells [1]. Horizontal wells are high-angle wells drilled to enhance the production from the hydrocarbon reservoirs. To obtain the maximum production in non-vertical wells, one has to keep drilling in the reservoir "Pay-Zone"; the largest production region inside the reservoir. In the process of drilling a borehole, geo-steering is the technique of accommodating the borehole position (inclination and azimuth angles) on the fly to reach one or more geological targets. Geo-steering technology is implemented in different ways [1]:

- Depending mainly on a pre-drilling plan and follow the static geological model without any change (Passive Geo-steering).
- Constructing the static model based on well drilling sensors measurements. In this technique, after drilling for a certain period the log data is processed, interpreted and the decision is made. (Reactive Geo-steering).
- Constructing the static model and adapts the drilling plan based on a certain depth interval. (Active Geo-steering).

The scope of this work is to develop a diminutive geological model from the offset pilot hole (the closest vertical well), and the lateral offset wells. The input data from the offset wells are: LWD (logging while drilling data “GR, Resistivity logs, porosity, and density”) [2], the trajectory (the survey sheet which contains the depth in MD, the inclination of the well at this depth, and the azimuth “the well direction” in the same depth), and the targets top depths. The main challenge of this phase is to build the a prediction for a sequence of the reservoir layers or the targets layers, which show varied dip angels and directions in the 3 dimensions. this geological problem input is summarized as the measurements of the drilling sensors inside the pre-drilled neighbor wells and the output is the predicted geological model that helps for drilling into a new well [3].

The problem of developing geological models can be stated as a machine learning problem and machine learning techniques can be applied. One of the these techniques is "Deep Learning and architectures". Deep Learning is broader family of Machine Learning techniques that tends to learn high level features from the given data. Thus, the problem it tackles is reducing task of making new feature extractor for each and every category of data (speech, image etc.). Deep Learning is inspired by the architectural depth of human brain, researchers tried for decades to train deep multi-layer neural network and they failed (with the exception of Convolution Network [4]). in 2006, Hinton [5] has presented Deep Belief Network (DBN) and in 2007, Bengio at el [6] proposed "Auto-encoders" for training deep architectures. DBN's are probabilistic generative model that generates both join and evidence probability. they contain many layers of hidden variables, each layer captures high-order correlations between the activities of hidden units in the layer below[5, 6]. DBN's are used for different machine learning problems, image understanding [5], Speech recognition [7], machine translation [8], etc..[9].

Why Not traditional shallow Back Propagation:

- It requires labeled training data while most data in this problem is unlabeled.
- It does not provide scalability for huge amount of training data.
- for noisy dataset, most probably it stuck into local optimal.

In this paper, a proposed multi-layers DBN is presented to construct a vision for the geological model inside a well that helps the geo-steering process by drawing an initial drilling path. The input for DBN's will be the logging sensors measurement for the surrounding wells and the output is the geological structure and proposed drilling path: inclination (Degree of drilling) and Azimuth (direction of drilling).

Section 2 illustrates the State-of-Art for geo-steering problem while Section 3 illustrates the deep topologies in general and DBN in specific. Section 4 presents the proposed system architecture section 5 shows the dataset and experimental results. Section 6 draws the conclusion and a future vision to scale the system capabilities.

2. Motivation and State of Art

Geoscientists find resources by assessing the characteristics and of the earth subsurface, which is called Geological model. The subsurface has been formed over thousands of decades and by the interaction of a host of sedimentary processes and time-varying boundary conditions like climate, sea level and tectonics [10]. The fundamental objective of geological characterization is predicting the spatial variation of geological variables. These geological variables can be any property of the geological subsurface that exhibits spatial variation and can be measured in terms of real numerical values .While the term "Reservoir" means a porous rocks or geological units hold the hydrocarbon reserves, the term "Pay Zone" means the section of the reservoir that contributes for real production [2]. Recently, the question turns from finding the reservoir to how to drill inside the pay zone to maximize its production. Also, considering all the deformations which causes the geological changes while drilling are a critical issue such as **"Faults".** To discover the presence of faults or folding during drilling requires an instantaneous decision to proceed in which direction. The guidance during drilling is the process of supporting the drilling with the future inclination and level to reach and keep drilling inside pay-zone [2]. Figure 1 shows the pre-drill well proposal plan path, the reservoir and pay zone model based on the already drilled neighbor wells data before drilling the well are illustrated.

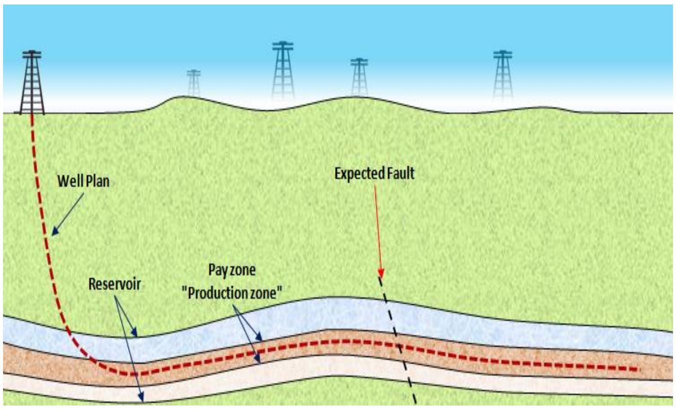

Figure 1: Pre-drill plan b ased on the surrounding wells.

The essential question that geologist has to answer is to calculate the "dipping angle" of the pay zone (β) which can be calculated as [1,2]:

$$X = \frac{A\tan\frac{y}{z}\times 180}{\pi} \quad (1)$$

$$\text{Where} \quad y/z = \frac{HD + (2\times DI)}{O\times 12} \quad (2)$$

$$\beta = (\text{theta} + X) - (90) \quad (3)$$

Where HD = hole diameter, DI = tools depth of investigation, O= offset depth of the image log and theta is the well inclination at the cut point. Planning an oil field exploration consists in assessing sets of wells to maximize oil recovery. Oil companies' products employ reservoir numerical simulation to new plan oil well exploration from already drilled wells. Traditionally, finite difference simulators are exploited, representing the model by a "discrete 3D mesh" of hexahedral cells [7]. The numerical simulator outputs a large set of data representing cell properties along a simulation time line. Among the generated data are 3D vector fields, especially flow of oil, gas, and water that must be analyzed by means of computer graphics visualization techniques[10, 11], Fig 2.

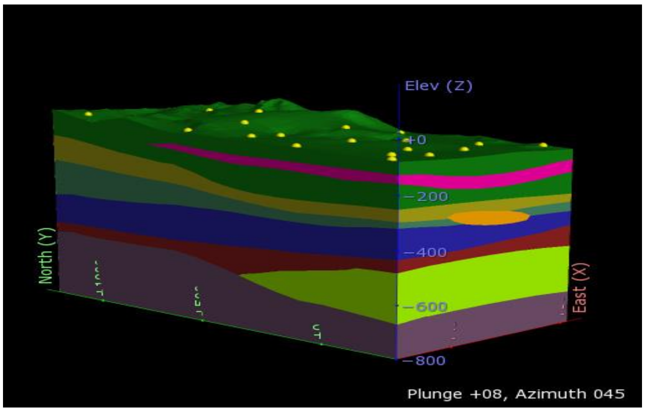

Figure 2: Visualized Geological model for different wells in the same region [2]

Researchers in Artificial Intelligence have implemented numerous quantitative approaches for the purposes of modeling the geological structure [10]. One of these approaches is traditional neural network. Schmoker et al [12] used feed forward network for computing Total Organic Carbon (TOC), Also Passey et al [13] built a neural network that maps between resistivity log, porosity logs and the rock type. Wong [14] exploited traditional neural network to compute the porosity of oil bearing rock. The developed system is capable of classifying rock quality for producing oil. Ashok at al[15] achieved 90% accuracy for classifying specific rock types. They built a vision system based on probabilistic neural network that could classify among 10 different types of rocks.

Rangwala [16] implemented Euclidian distance measures and K-nearest Neighbors (KNN) to extract and classify the rock features. Preston et al. [9] developed a multivariate analysis system that extracts features from the acoustic signals and unsupervised clustering module to recognize the rock type.

The main objective of this research is to propose the drilling plan based on constructed geological model for the surrounding (n) wellbores. Figure 3 illustrates the conceptual design for the proposed system.

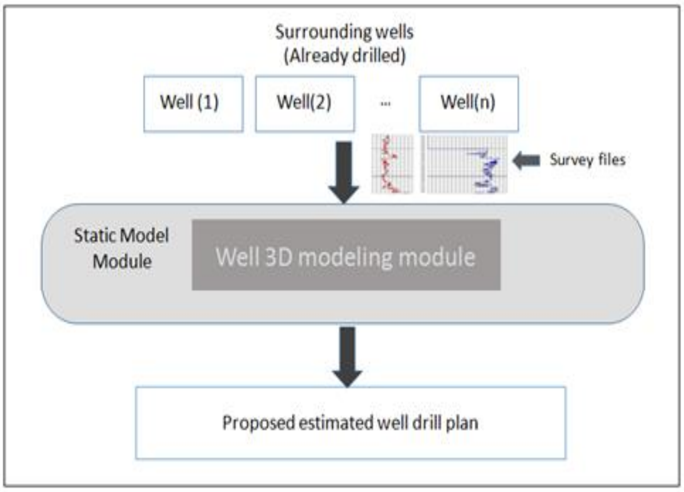

Figure 3: Proposed system

Due to the wide range of output dip angles possible values, and the remarkable variations for the possibility of existing fault, using traditional AI techniques such as Fuzzy logic, Case-Based Reasoning (CBR) is non-practical. Also, the employment of discriminative techniques such as traditional neural network requires a huge amount of training data. This paper exploits DBN with a supervised output layer and fine-tuned with back-propagation algorithm as generative technique for predicting the geological model. The proposed DBN solution outperformed traditional neural network, multiplicative neural network and CBR for both limited and large dataset.

3. Deep Learning

A standard artificial neural network (ANN) consists of many simple, connected processing elements called neurons, each producing a sequence of real-valued activations. Input layer neurons get internally activated through perceiving the environment values, other layers get activated through weighted connections from previously neurons. An efficient gradient descent method for teacher-based Supervised Learning in discrete, differentiable networks of arbitrary depth called back propagation (BP) was developed in the 60s and 70s. BP-based training of deep NNs with many layers, however had been found to be inefficient and unstable in practice. DL became practically feasible to some extent through the help of Unsupervised Learning. Recently, deep NNs have finally gained wide-spread attention, mainly by outperforming other machine learning techniques in numerous applications. Since 2009, supervised deep NNs have won many world-wide pattern recognition competitions [17], achieving the first superhuman visual pattern recognition results in limited domains. Deep NNs also have become relevant for the more general field of Reinforcement Learning (RL) where the supervising teacher is absent. One major factor that facilitates the application of DL in practical life is the availability of GPUs technology [17].GPU technology permits DL implementation for video understanding, image understanding, speech recognition, machine translation, etc...

One of deep architectures that is commonly used for extracting features in multiple levels is DBN. Deep Belief Networks (DBNs) consists of a number of layers of Restricted Boltzmann Machines (RBMs) which are trained in a greedy layer wise fashion. A RBM is a generative undirected graphical model, Fig 4 [8].

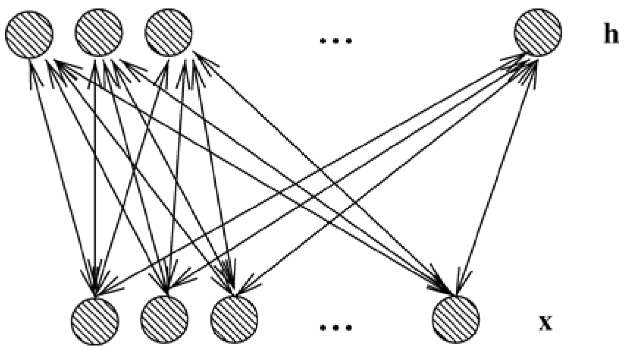

Figure 4: DBN topology

The lower layer (x) is the visible layer (for this layer) and the top layer h as the hidden layer. The visible (input) and hidden layer units are stochastic binary variables. The weights between the visible and the hidden layer are undirected and are denoted as (W). In addition, each neuron has a bias parameter. The model presents the probability distribution [8]:

$$P(x,h) = \frac{e^{-E(x,h)}}{Z} \quad (4)$$

With the energy function, E(x, h) and the partition function Z is defined as:

$$E(x,h) = -b'x - c'h - h'Wx \quad (5)$$

$$Z(x,h) = \sum_{x,h} e^{-E(x,h)} \quad (6)$$

Where b and c are the biases of the visible layer and the hidden layer respectively. The sum over x, h represents all possible states of the model. The conditional probability of one layer, given the other is [8]:

$$P(h|x) = \prod_i P(h_i|x) \quad (7)$$

if one layer is given, the distribution of the other layer is factorial. Since the neurons are binary the probability of a single neuron being on is given by:

$$P(h_i = 1|x) = \frac{\exp(c_i + W_i x)}{1 + \exp(c_i + W_i x)}$$
$$P(h_i = 1|x) = \mathrm{sigm}(c_i + W_i x) \quad (8)$$

Similarly the conditional probability for the visible layer can be computed as:

$$P(x_i = 1|h) = \mathrm{sigm}(b_i + W_i h) \quad (9)$$

It could be viewed as a probabilistic version of the normal sigmoid neuron activation function. To train the model, the idea is to make the model generate data like the training data. From Mathematical point of view we look for maximizing the log probability of the training data which is equivalent to minimize the negative log probability of the training data.

4. Proposed system architecture

The main target is to determine the drilling parameters: Inclination and Azimuth for each drilling step. To achieve this goal, intermediate steps and conclusions must be accomplished, Fig 5 illustrates the proposed system with the intermediate hidden layers.

- An input layer that processes the drilled surrounding wells measurements
- A hidden layer that extracts the correlation among the measurements in the surrounding .
- A hidden layer that is trained to predicts the faults and recognizes the faults length, direction and type.
- An output layer that is supervised trained to predict the next drilling step parameters

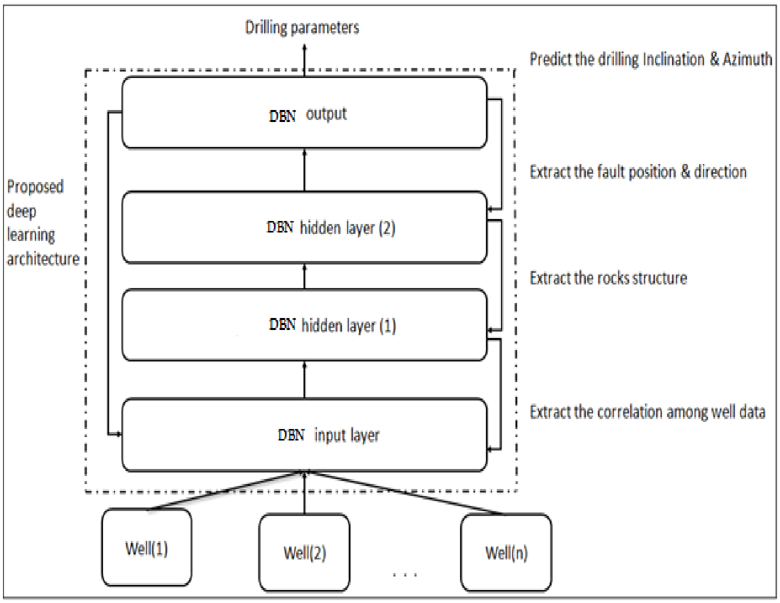

Figure 5: The proposed system architecture.

5. Dataset and Results

This paper addresses the problem of constructing a regional geological model for the purpose of developing a drilling plan for a new wellbore based on surrounding wells. A.S. Elons et al [18] proposed a PSO-based multiplicative NN for the same objective and reached 88% accuracy for the extracted plan compared with real drilling plan. In this paper, Multi-layer DBN system is proposed and compared with results also, the proposed system is compared with a Case-Based Reasoning (CBR) and simple Expert System these were developed in house to address the same issue. The experiments were conducted on 3 groups of wells, each group consists of 5 wells (4 used for input and 1 is used as a desired). 2 groups were used for training purpose and the remainder group is used for testing.

The input files are M/LWD files [2] which consist of:

- Inclination and Azimuth against the True Vertical Depth (TVD).
- Resistivity, Porosity, Density and Gamma Ray (GR).

Table 1 illustrates the accuracy for the test well for the proposed technique (Deep Learning) against other techniques: CBR, Expert System, Traditional Neural Network and Multiplicative Neural Network against the number of wells.

Table 1: Comparative study between the proposed system and in-house built other techniques

Data Group	Number of input wells	Drilling plan Accuracy %				
		Proposed DBN	CBR	Expert System	Neural Network	Multiplicative NN
1	1	36	17	71	33	32
	2	47	19	73	35	39
	3	58	33	72	47	47
	4	58	36	72	55	58
	5	60	44	74	58	59
2	6	64	54	64	58	60
	7	71	59	65	57	62
	8	77	59	60	64	65
	9	89	59	55	68	68
	10	88	64	54	67	69
3	11	90	63	53	73	72
	12	90	63	55	76	80
	13	92	60	52	71	81
	14	**92**	61	48	68	79

Expert system started with a very promising results for limited training dataset but the conflicts among rules in the knowledge base degrade its accuracy. CBR system also suffers from accuracy due to similarity criteria choice. Neural network shows better results as the training data increases but up to limit then the problem of over-fitting affects the results. Proposed Deep Learning system performance increases as the available training data is used. Figure 6 illustrates ROC curve of error analysis for DBN system

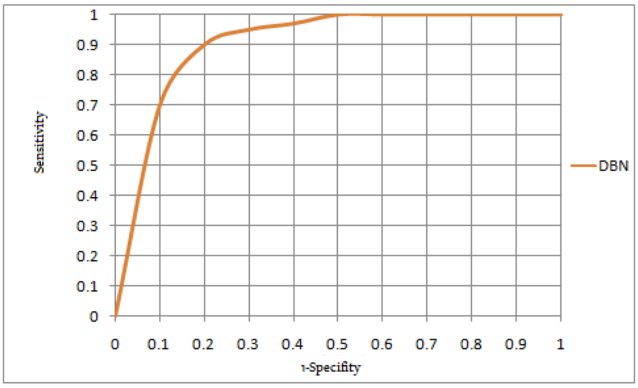

Figure 6: ROC Curve

The input layer consists of 7 units and the output consists of 180 units: 90 units for inclination angles in range of 0-90 degrees and 90 units for the Azimuth. Hidden units in the first hidden layer consists of 120 units, the second consists of 90 and the third consists of 60 units. With the layer-wise training algorithm for DBNs, one element that we would like to dispense with is having to pick the number of training iterations for each layer. It would be efficient if we did not have to explicitly add layers one at a time. the system is trained as all layers simultaneously, but keeping the "greedy" idea that each layer is pre-trained to model its input, ignoring the effect of higher layers. This idea can be implemented by reaching an accepted but optimal error in each layer and then train all the layers simultaneously. Training of the DBNs was accelerated by exploiting a graphics processor unit (GPU). A single pass over the entire training set (14 wells) during pre-training consumes around 3-4 minutes per layer. An epoch of fine-tuning with back-propagation took around 10 and 30minutes with frame-based, and sequence based criteria, respectively.

6. Conclusion

This paper addresses the problem of constructing a geological model for the purpose of proposing a drilling path inside a new wellbore based on surrounding wells. System input is survey files contain sensors measurements inside surrounding wells against both vertical and inclined depths. Different AI techniques were explored, Expert system, Traditional and Multiplicative Neural Network, CBR and Deep learning. Deep Belief Network shoed superiority over other techniques and reached 92% accuracy from real drilling happened. 15 wells survey files were exploited and a comparative study has been conducted.

7. Acknowledgement

This work is a part of a research project funded by " Information Technology Academia Collaboration" (ITAC) in Egypt. Also we would like to thank Dr. Ahmed Samir. Ain-Shams University for his consultation

References

[1]. D. Machado , F. Solano, A. Roberto deMedeiros, M. Rodrigues and F. de Azevedo,2013, "Pipeline Shore Approach Installation by Horizontal Directional Drilling, ASME 2013 32nd International Conference on Ocean, Offshore and Arctic Engineering Volume 4B: Pipeline and Riser Technology Nantes, ,(June 9–14, 2013), France

[2]. "Oilfield Glossary". Retrieved in Oct 29, 2015.

[3]. Quinn, T.H., J. Dwyer, C. Wolfe, S. Morris, and D. Coope, 2008, Formation Evaluation Logging While Drilling (LWD) in Unconventional Reservoirs for Production Optimization, SPE Eastern Regional/AAPG Eastern Section Joint Meeting, Pittsburgh, Pennsylvania, SPE no.119227, USA

[4]. Y. LeCun, K. Kavukcuoglu, and C. Farabet, 2010, Convolutional networks and applications in vision. In Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on, 253–256. IEEE.

[5]. Hinton, G., Osindero, S., and Teh, Y.,2006. A fast learning algorithm for deep belief nets, Neural Computation, vol. 18, 1527-1554.

[6]. Hugo Larochelle, YoshuaBengio, Jerome Louradour, and Pascal Lamblin., Exploring strategies for training deep neural networks, 2009. The Journal of Machine Learning Research, 10, 1–40.

[7]. A. Mohamed, D. Yu, and L. Deng, Investigation of Full-Sequence Training of Deep Belief Networks for Speech Recognition, 2010, in Proceeding . Interspeech.

[8]. R. Sarikaya, G. Hinton and B. Ramabhadran, 2011, Deep Belief Networks for Natural Language Call-Routing ", In Proceeding. ICASSP.

[9]. J. Zhang, D. Zhang, and J. Hao, 2015, Local Translation Prediction with Global Sentence Representation, to be published in Proceeding . Int'l J. Conf. Artificial Intelligence.

[10]. Rasmus, J.C., C. Esmersoy, J. Seydoux, and A. Hawthorn, 2005, LWD for Imaging, Wellbore Placement, and Formation Evaluation, Offshore Technology Conference, Houston, Texas, USA, no. 17646,

[11]. Wang, T. and J. Signorelli, 2004, Finite-difference modeling of electromagnetic tool response for logging while drilling, Geophysics, 69, 152–160..

[12]. Schmoker, J., and Hester, T., 1983, Organic Carbon in Bakken Formation, United States Portion of Williston Basin: American Association of Petroleum Geologists Bulletin, 67,2165- 2174

[13]. Passey, Q., Creaney, J., Kulla, F., Moretti, F.and Stroud, J., 1990, A practical model for organic richness from porosity and resistivity logs, AAPG Bull., Volume. 74,1777-1794.

[14]. Wong, Jian F.X. and Taggart I.J. 1995,A critical comparison of neural networks and discriminate analysis in lithofacies, porosity and permeability predictions. Journal of Petroleum Geology 18, 191-206.

[15]. A.Kumarand S. Chatterjee,2014, Computer vision-based limestone rock-type classification using probabilistic neural network , Geoscience Frontiers, 1-9.

[16]. S. Rangwala and D. Dornfeld, 1990, Sensor Integration Using Neural Networks for Intelligent Tool Condition Monitoring,Journal of Engineering for Industry, Volume. 112, 219.

[17]. K. Cho et al., 2014, Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation, Proceeding. Empirical Methods and Natural Language Processing, 355–362.

[18]. A.S.Elons, D. Ahmed, M. Elgendy, 2016, A Proposed Model for Predicting the Drilling PathBased on Hybrid PSO-BP Neural Network, to be published in proceeding of SAI Computing Conference, London.