Iranian Journal of Oil & Gas Science and Technology, Vol. 10 (2021), No. 3, pp. 83–98
http://ijogst.put.ac.ir

Estimation of Drilling Mud Weight for Iranian Wells Using Deep-Learning Techniques

Aref Khazaei[1], Reza Radfar[2*], and Abbas Toloie Eshlaghy[3]

[1] Ph.D. Candidate, Industrial Management, Science and Research Branch, Islamic Azad University, Tehran, Iran.
[2] Professor, Industrial Management, Science and Research Branch, Islamic Azad University, Tehran, Iran.
[3] Professor, Industrial Management, Islamic Azad University, Science and Research Branch, Tehran, Iran.

Highlights

- Collecting an extensive data set with various features of wells obtained by merging the data of the United Kingdom, Norway, and Iran fields;
- Presenting an accurate mud weight estimator using a deep-learning model;
- Challenging the presented model in real-world conditions and providing solutions to improving its performance.

Received: July16, 2021; *revised:* August 02, 2021; *accepted:* August 07, 2021

Abstract

Iran is one of the largest oil and gas producers in the world. Intelligent manufacturing approaches can lead to better performance and lower costs of the well drilling process. One of the most critical issues during the drilling operation is the wellbore stability. Instability of wellbore can occur at different stages of a well life and inflict heavy financial and time damage on companies. A controllable factor can prevent these damages by selecting a proper drilling mud weight. This research presents a drilling mud weight estimator for Iranian wells using deep-learning techniques. Our Iranian data set only contains 900 samples, but efficient deep-learning models usually need large amounts of data to obtain acceptable performance. Therefore, the samples of two data sets related to the United Kingdom and Norway fields are also used to extend our data set. Our final data set has contained more than half-million samples that have been compiled from 132 wells of three fields. Our presented mud weight estimator is an artificial neural network with 5 hidden layers and 256 nodes in each layer that can estimate the mud weight for new wells and depths with the mean absolute error (MAE) of smaller than ±0.039 pound per gallon (ppg). In this research, the presented model is challenged in real-world conditions, and the results show that our model can be reliable and efficient in the real world.

Keywords: Artificial Neural Networks, Deep learning, Drilling Mud Weight, Mean Absolute Error, Smart Manufacturing

[*] Corresponding author:
Email: r.radfar@srbiau.ac.ir

How to cite this article
Khazaei, A., Radfar, R., Toloie Eshlaghy, A., Estimation of Drilling Mud Weight for Iranian Wells Using Deep-Learning Techniques, Iran J. Oil Gas Sci. Technol., Vol. 10, No. 3, pp. 83–98, 2021.
*DOI: http://dx.doi.org/10.22050/ijogst.2021.138844,**This is an Open Access article under Creative Commons Attribution 4.0 International License.(creativecommons.org/licenses/by/4.0) /*

1. Introduction

Smart manufacturing is one of the main sectors of modern industry. It usually applies new data analytics and machine-learning methods. The main goals of intelligent manufacturing usually are minimizing the human roles and presenting a self-management system. Successful systems usually reduce production costs and waste time, improve quality, and increase resource productivity (De Felice et al., 2018). This research attempts to smarten the mud weight estimation.

In the oil and gas industry, new artificial intelligence technologies can lead to better performance and lower costs of the well drilling process (Yavari et al., 2018; Sebtosheikh et al., 2015). The wellbore stability is one of the most critical issues during the drilling operation. Instability of wellbore can interrupt the drilling operation and can waste time and money; even in worse cases, it can endanger production from several other wells (Zeynali, 2012). By upsetting the balance between the load of the rocks around the wellbore and the stress of concentrations during drilling, two main problems may occur for the stability of the wellbore: breakout and fracture (Peng and Zhang, 2007). Breakout is equivalent to formation collapse. Borehole breakout failure may lead to fluid flow into the well, called kick (Abbas et al., 2019). A fracture is any separation in a geologic formation, such as a joint or a fault that divides the rock into two or more pieces (Park, 1989). The breakout of the wellbore, the entrance of unexpected fluids into it, hanging up of drill string, appearance of cracks and fractures in the wellbore, and consequently the loss of drilling fluid into the formation are some of the harmful consequences of wellbore instability; however, most of them can be prevented by choosing the proper mud weight (McLellan, 1996). A stable wellbore should have an appropriate balance between the uncontrollable factors of earth stresses, rock strength, pore pressure, and the controllable factors such as drilling mud weight (Cheatham, 2012). In the drilling industry, the drilling mud is used to keep up the stability of the wellbore; its weight is an essential factor to take care of the hydrostatic pressure and stability of the well (Peng and Zhang, 2007).

Wellbore instability can occur at different stages of a well life, such as drilling, well completion, and production (Peng and Zhang, 2007). The financial damages of instability are estimated between 500 million to 1 billion dollars per year, depending on the severity of the problem (Zeynali, 2012). Hence, the mud weight estimation and determination of the equivalent circulating density (ECD), i.e., the dynamic density affected by friction, are the essential controllable factors. A proper estimation can significantly save time and money (Peng and Zhang, 2007). Therefore, many researchers have tried to help oil industry managers by presenting mathematical, experimental, and artificial intelligence (AI)-based mud weight estimation models (e.g., Aadnoy and Chenevert, 1987; Zoback et al., 2003; Zhou et al., 2016).

Artificial intelligence, a branch of computer science, encompasses all processes that lead to human or animal intelligence simulation on computers or computer-controlled devices. Written sources and AI researchers describe this field as “the study and design of intelligent agent” in which an intelligent element or agent is a system that understands the environment and takes steps to maximize its chances of success (Kim, 2017). More than 80% of the companies that use AI technologies could increase their productivity up to 17%–20% (Tuptuk and Hailes, 2018). Further, in the oil and gas research area, the researchers try to increase the efficiency of financial and human resources using AI methods. They have

used various AI methods for simulation, prediction, estimation, optimization of different issues in the oil and gas industry.

The primary goal of this research is to estimate the drilling mud weight for Iranian wells using deep-learning techniques. Deep learning is one of the most popular and practical AI techniques. Only 900 samples of Iranian wells have been provided by Pars Drilling Fluids (PDF) Company to build the estimator model (PDF, 2021). More samples are needed to obtain a more accurate model using deep-learning techniques. Therefore, the samples of two data sets related to the United Kingdom and Norway fields are also used to build the model. The obtained model of our research can assist Iranian engineers in choosing safe mud weight in operational conditions and reduce the human role and error by creating a self-management system. The main contributions of our research are:

- Collecting an extensive data set with various features of wells obtained from merging the data of the United Kingdom, Norway, and Iran fields;
- Presenting an accurate mud weight estimator using a deep-learning model.

This paper is organized as follows: Section 2 presents the relevant literature. In Section 3, our collected data set is introduced, and Section 4 describes our methodology. In Section 5, different experiments and their obtained results are discussed, and finally, Section 6 concludes the paper.

2. Literature review

Many studies have been done on the stability of the wellbore and the factors affecting it and estimating the proper mud weight to achieve a stable well. Aadnoy and Chenevert (1987) were two of the first researchers that used rock mechanics to analyze the wellbore stability. In this study, using some rock mechanics features, the stresses around the wellbore that affect the stability of the well were calculated (Aadnoy and Chenevert, 1987).

Aadnoy and Belayneh (2004) proposed an elastoplastic fracturing model for wellbore stability analysis. Hareland and Dehkordi (2007) and Bakhtiarizadeh et al. (2009) also studied the wellbore stability and determined the proper mud weight for the wellbore stability through the obtained models.

Some researchers, such as Zoback et al. (2003), Alsiyabi et al. (2019), and Kamgue Lenwoue et al. (2019), tried to study and analyze the wellbore stability through the typical failure criteria, including Mohr–Coulomb, Mogi, and Mogi–Coulomb and provided mud weight estimators using these standard failure criteria. Han et al. (2014) used numerical simulations to select the safe mud weight window.

In a nutshell, some oil and gas researchers (e.g., Liu et al. (2018), Khatibi et al. (2018), Deng et al. (2016), Silva et al. (2017), and Shen et al. (2014)) studied and presented mathematical and experimental models to predict the proper mud weight.

The approximately permissible drilling mud is placed between pore pressure (or formation pressure) and fracture pressure. These upper and lower bounds determine the acceptable mud weight window (Pattillo, 2018). There are several ways to calculate these two pressures, including pressure evaluation logs and formulas. For example, in claystone/shales lithology, the Terzaghi principle links the vertical stress (S_v) to the pore pressure (P_p) and the effective vertical stress (σ'_v) (Aird, 2019):

$$P_p = S_v - \sigma'_v \tag{1}$$

The relationship between rock fracture pressure (P_f) (psi) and pore pressure (P_p) (psi) is expressed by (Meyers, 2016):

$$P_f = (\sigma_{ob} - P_p)\left(\frac{v}{1-v}\right) + P_p \tag{2}$$

where σ_{ob} is overburden stress (psi), and v indicates Poisson's ratio.

In addition to mathematical and experimental studies, some of which are mentioned above, in recent years, some researchers have turned to artificial intelligence deep-learning techniques in the field of the oil and gas industry. Bandura et al. (2018) used deep learning to provide methods for interpreting seismic data. Another deep-learning-based research done by Li (2018) aimed to automatically detect and classify geological structure elements (such as salt domes faults) based on seismic images. Li et al. (2019) proposed a deep-learning estimator to predict long-term well performance based on monitoring data.

In summary, in recent years, artificial neural networks and deep-learning techniques as a subset of machine learning have been used to solve various oil and gas problems. Some of the studies on estimating and predicting the amount of lost circulation, fracture pressure, drilling stuck pipe, and the rate of penetration include Behnoud Far and Hosseini (2017), Ahmed et al. (2019), Shadizadeh et al. (2010), and Bataee et al. (2014).

Some other studies used artificial intelligence methods to estimate proper mud weight. Zahiri et al. (2019) presented a mechanical earth model based on well logging data, elastic moduli of rock, and failure criteria. They also applied artificial intelligence and machine learning algorithms to find a relationship between well logging data and the safe mud weight window. Abbas et al. (2020) conducted a similar study and used machine learning algorithms, including artificial neural networks, to estimate the mud weight.

The mud weight is an important parameter, especially when drilling in salt zones. In such cases, it is usually determined by expert designers and is updated during the drilling operation. Petrobras Company presented an application to estimate the proper mud weight to prevent salt fluency in susceptible zones using lithology data and history of wells drilled in pre-salt zones. Petrobras Company applied machine learning techniques (Pereira et al., 2013).

Drilling at great depths of more than 7000 m has always been challenging because of high pressures and temperatures. On the other hand, the density of drilling fluid is usually affected by this high temperature and pressure. Zhou et al. (2016) attempted to model the drilling fluid density for such situations. This model was presented using artificial neural networks.

Phan et al. (2020) used artificial intelligence to present a time-dependent mud weight window predictor. They also displayed polar charts for the predicted mud weight in inclined wellbores and investigated the effect of neural network architecture on the model error.

Table 1

Comparing the most relevant research on predicting mud weight.

Author (s)	Method (s)	Sample size	Input Parameters	Results
(Zahiri et al., 2019)	AI and ML algorithms (Ensemble Regressors, Ridge Regression, SVM–RBF)	Three wells	Well logging data	**For SVM–RBF method MSE = 4.36 MPa**

Author (s)	Method (s)	Sample size	Input Parameters	Results
(Pereira et al., 2013)	Random forest	1047 samples (40 wells)	Depth, rock type, and temperature	**RMSE = 0.38 ppg**
(Zhou et al., 2016)	Artificial neural network (One hidden layer)	120 samples	Oil phase volume fraction, water phase volume fraction, temperature difference, pressure difference, and initial density	**MAE = 0.005 sg**
(Phan et al., 2020)	Linear regression, decision tree, extra random forest, artificial neural network	About 2.5 million samples have been generated randomly.	True vertical depth, inclination, azimuth, overburden stress gradient, max and min horizontal stress gradient, pore pressure, max horizontal stress azimuth, Young's modulus, Poisson's ratio, cohesion, friction angle, tensile strength, permeability, Biot's coefficient, Skempton's coefficient, radial ratio, and time	**For the neural network method, MAE = 0.33 ppg**
(Tewari, 2019)	Random, forest, Bagging, SVR, and artificial neural network	5000 samples (14 wells)	Measured depth, true vertical depth, rate of penetration, weight on bit, torque, rounds per minute, standpipe pressure, flow rate, total gas, mud weight, D-exponent, lithofacies types, inclination, azimuth, and pore pressure	**For the ANN method, MAE = 0.0445 kg/l**

Tewari (2019) used several machine learning methods in a study to estimate the safe mud weight. This research data set was collected from a Norwegian field, and the different obtained models were compared together.

In summary, Table 1 compares the most relevant research with ours with more details. These studies focused on mud weight estimation using different artificial intelligence methods.

3. Research data set

In deep-learning modeling, to obtain an appropriate and accurate model the results of which can be extended to new samples and can be used in operational conditions, a data set with accurate and enough samples is needed. As mentioned before, only 900 samples of Iranian wells have been provided by Pars Drilling Fluids (PDF) Company (PDF, 2021). This number of samples is not enough, and more samples are needed to obtain a more reliable model using deep-learning techniques. Therefore, the samples of two data sets related to the United Kingdom and Norway fields are also used to build the model.

Therefore, our research data set is collected from three data sets related to some oil and gas fields in the United Kingdom (NDR, 2021), Norway (Equinor-Company, 2021), and Iran (PDF, 2021). National Data Repository (NDR) has published the UK Oil and Gas Authority (OGA) data set. This data set contains the information on wells from exploration to the shutdown of offshore wells drilled by subsidiary companies. In the OGA data set, well logging reports, such as drilling evaluation log, pressure evaluation log, and formation evaluation log; digital records; and some other documents collected over five decades of operation on the UK Continental Shelf (UKCS) were recorded (OGA,

2021). The second data set of our research is Volve. This field is located 200 km west of Stavanger, in the south of Norway, at a depth of 80 m below sea level, which was discovered in 1993, and its production began in 2008. Volve data set is available to researchers by Equinor Company, under a modified Creative Commons license and Equino Open Data license (Equinor-Company, 2021; Norwegian-Petroleum, 2021). The third data set of our research has been provided by PDF Company (PDF, 2021). The Iranian field is located in the southwest of Iran (the west of the Karun river) and is a border field between Iran and Iraq.

OGA, Volve, and PDF data sets have been processed and integrated to collect our final data set. The data samples are extracted from well-logging curves and drilling reports, and they have features with two following characteristics: 1) There is a logical relationship between the feature and mud weight; 2) This feature has been recorded in all used data sets. Each data set sample is presented by 13 features and its corresponding mud weight value. Table 2 contains more details related to the features. These features are used as inputs of the mud weight estimator.

Table 2

The list of final features in our data set.

Variables	Units	Minimum	Maximum
Measured Depth (MD)	m	33.00	6361
Hole Diameter	Inches	3.75	36
Inclination	Degree	0.00	90
Azimuth	Degree	0.00	359
Rate of Penetration (ROP)	m/h	0.03	724
Weight on Bit (WOB)	klb_f	0.00	83
Rounds Per Minute (RPM)	rpm	0.00	815
Torque (TQ)	$klb_f.ft$	0.00	44
Stand Pipe Pressure (SPP)	psi	4.00	7527
Mud Flow Rate	USgal/min	0.00	1450
Mud Temperature Out	°F	10.80	407
Total Gas(TG)	%	0.00	72
Lithology	Categorical	0.00	15

Figure 1 shows some scatter plots related to the input features and target feature of our final data sets (drilling mud weight).

More than half a million samples were obtained after integrating OGA, Volve, and PDF data sets. Some of these samples have at least one Kick, Gain, and Lost Circulation problem. One of the main reasons for these problems during drilling is improper mud weight. In this research, these problematic samples are ignored to model. Because of the unsuitable mud weights, using these samples during the modeling can mislead the model. In other words, the model is trained only with the correct samples and weights that have done their task correctly. After ignoring these samples, the number of remaining samples is 477270.

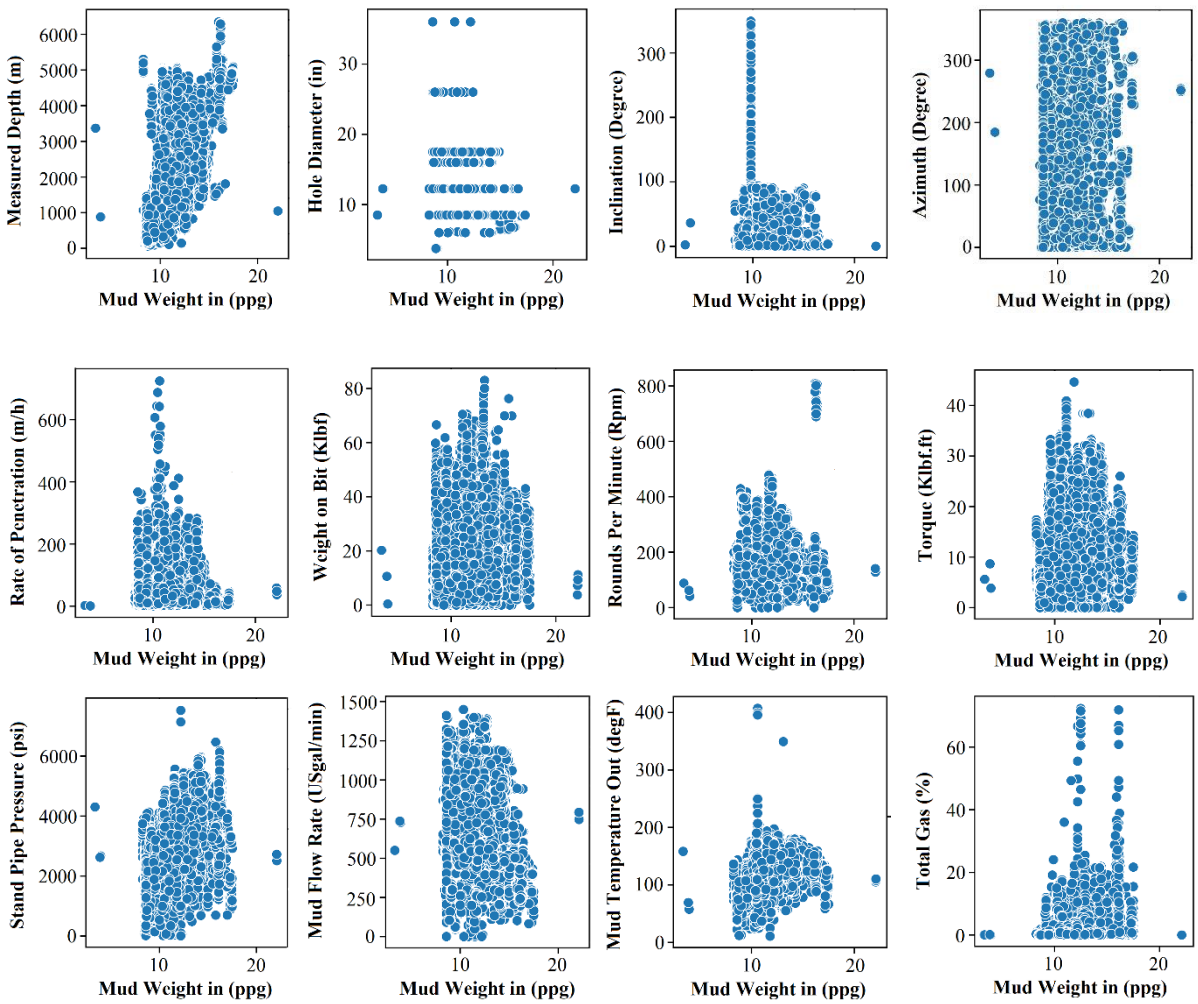

Figure 1
Relationships of input and output features.

3.1. Data preprocessing

After data collection, the next step is data preprocessing. The raw data usually need some processes to become ready for modeling. Data preprocessing has four significant tasks generally, namely data integration, data cleaning, data reduction, and data transformation, all or some of which may be necessary to prepare the data (García et al., 2015). The following tasks were performed on our data set as preprocessing:

- Ignoring or estimating missing values. Missing values refer to values that have not been recorded in the data set for any reason. In this research, samples with at least 10 missing values for consecutive depths have been ignored. The other samples with missing values have been estimated using the mean, median, or mode of their neighborhood samples (depending on the type of missed features).
- Ignoring the outlier samples. An outlier is a sample that is much different from the others. The samples that at least one (or more than one) of their features have uncommon and outlier (too high or low) values have been identified and ignored. For example, the samples with a recorded mud temperature of above 600 °F were ignored. The occurrence of these temperatures during the drilling is not possible.
- Normalizing the numerical features;

- Integrating the samples of three data sets;
- Converting the lithology (a categorical feature) to some binary features through the One Hot encoding method.

After data preprocessing, in our data set, there are 414170 samples from 91 wells on the UK Continental Shelf, OGA data set (86.7% of the total data); 62885 samples from 25 wells in the Volve field of Norway (13.2% of the total data); and 215 samples from 16 wells in the southwest of Iran (0.045 % of the total data). These data are now ready to build our mud weight estimator model using deep-learning techniques.

4. Methodology

This research aims to present a mud weight estimator for Iranian wells using deep-learning techniques. Deep learning is a specific subset of machine learning, and a deep model includes multiple layers. After completing the learning phase, each layer has a more meaningful representation of the data than the previous layer. Most of the time, in deep-learning models, the representational layers are artificial neural layers. Artificial neural networks include the input layer, the output layer, and one or more hidden layers. In other words, a deep-learning model is usually an artificial neural network with several hidden layers, including several nodes. The training phase (the phase of learning) is adjusting the weights of the network neurons based on the input samples (Chollet, 2017).

Python is one of the most popular programming languages in data science, which includes many libraries for scientific calculations and machine learning (Raschka and Mirjalili, 2017). Keras is one of these libraries that is so efficient and popular in deep learning (Chollet, 2017). Python and Keras have implemented our mud weight, estimator model.

The main pre-requirements for building a model are (Chollet, 2017):

- Input data samples: In this research, they are the samples represented by 13 features;
- Expected outputs: In this research, it is the drilling mud weight (the density of drilling fluid used in operational conditions);
- Evaluation measure of the model (also called the cost function): Mean absolute error (MAE) in this research. This metric evaluates the difference between the actual mud weight and the estimated mud weight by the model.

  $$\text{MAE} = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n}$$

  where y_i is actual value, $\hat{y}_i$ is estimated value, and n is the number of test samples. The lower MAE values indicate that model works better.

Regression analysis is a subfield of supervised machine learning. The regression goals are to model the relationship between features and a continuous target feature. In our research, the continuous target feature is the mud weight. In a nutshell, to present a mud weight estimator, the following steps are performed:

- Through the activation function, the input data are processed to achieve their corresponding outputs;
- The cost function compares the outputs of the model with the actual outputs, and the optimizer function updates the current neural weights;

As mentioned in Section 3, our model is presented using 477270 final samples. During the modeling, 80% of samples are used to train, and the remaining 20% of samples are employed to test the performance of the designed model.

This study attempted to find the best values for hyper-parameters and model structure by trials and errors. After trying more than 200 different models, a model with the following characteristics have been selected:

- A seven-layer neural network with five hidden layers, an input layer, an output layer;
- Each layer has 256 nodes;
- The activation function of each node is Rectified Linear Unit (RELU);
- The model optimization algorithm is Adam, with a learning rate of 0.001. The learning rate is the amount that model moves toward the goal (optimal value) in each step;
- The weight updating method is mini-batch, and the batch size is 128;
- The epoch number (the number of training repetitions) is 100 times.

Figure 2 shows the summary of our model. As mentioned before, the first layer is a normalizing layer that normalizes the input samples, and the following five layers are hidden layers. The last layer is the output layer that provides the estimated mud weight as the model output. All the layers are dense (fully connected). In a dense layer, each neuron receives input from all neurons of its previous layer, and the dense layer is the most commonly used in neural network models (Kim et al., 2019). The model is sequential, implying that in each layer, the entrance is from one side and the exit from the other side; in other words, the model consists of several linear layers (Chollet, 2015). The second column of Figure 2 is related to the shape of each layer output, and the third column shows the number of parameters per layer. The total number of the parameters is 270648, and 270593 of them are trainable; the remaining are related to the normalizing layer and untrainable.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
normalization (Normalization multiple                  55
_________________________________________________________________
dense (Dense)                multiple                  7168
_________________________________________________________________
dense_1 (Dense)              multiple                  65792
_________________________________________________________________
dense_2 (Dense)              multiple                  65792
_________________________________________________________________
dense_3 (Dense)              multiple                  65792
_________________________________________________________________
dense_4 (Dense)              multiple                  65792
_________________________________________________________________
dense_5 (Dense)              multiple                  257
=================================================================
Total params: 270,648
Trainable params: 270,593
Non-trainable params: 55
```


```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
normalization (Normalization multiple                  55
_________________________________________________________________
dense (Dense)                multiple                  7168
_________________________________________________________________
dense_1 (Dense)              multiple                  65792
_________________________________________________________________
dense_2 (Dense)              multiple                  65792
_________________________________________________________________
dense_3 (Dense)              multiple                  65792
_________________________________________________________________
dense_4 (Dense)              multiple                  65792
_________________________________________________________________
dense_5 (Dense)              multiple                  257
=================================================================
Total params: 270,648
Trainable params: 270,593
Non-trainable params: 55
```


Figure 2
The summary of the structure developed herein.

The generalization of a model is an expression used to describe a model's ability to predict/estimate the target features of new samples. If a model does not train enough (too little or too much), it cannot generalize what it learned (underfitting or overfitting). A validating data set has been used to avoid these situations. In our research, based on the rule of thumb 8:2 (Kim, 2017), the validating data set is 20% of the training data set. The validating data set aims to monitor the performance of the model during the training process that is used to estimate model performance while tuning the model's weights and hyper-parameters. In order to check the training process and avoid overfitting/underfitting, the graph of the mean absolute errors (MAEs) for the training and validating data sets during the training epochs has been plotted (Figure 3). There is no significant difference between the training and validating curves, so the phenomenon of overfitting has not happened.

In our experiments, on average, 64 min is spent on model training. As shown in Figure 3, the final MAEs for the training and validating data set is 0.041. The average MAE for the test data set is 0.03849, indicating that, on average, the estimated mud weight values by our presented model have a ±0.03849 (ppg) difference from the actual values.

In the next section, our presented model has been challenged using Iranian samples, and the various experiments have simulated the real-world conditions.

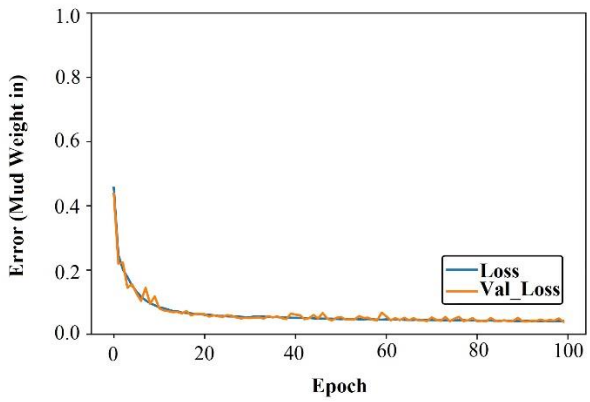

Figure 3
The MAE values for training and validating data sets during the training epochs.

5. Results and discussion

These experiments aim to challenge and evaluate the proposed model performance using one well of Iran. Therefore, the following steps have been done:

1- One well from the Iranian field was selected (202-107);
2- All the samples related to this selected well were isolated from our data set;
3- The model was built by using the remaining samples. The model structure and hyperparameter values were similar to the previous section. In this step, the MAE of the model was 0. 05224 ppg;
4- Then, the model was evaluated using the isolated well samples; the model had not used any of these samples previously. The MAE of the model for these test samples was 1.15461 ppg.

In Figure 4, for each depth, the actual mud weight values used during drilling operations are compared with the estimated mud weight values by the proposed model for this selected well. As shown in Figure 4, the model is not entirely successful in predicting the fluctuation of the mud weight in a depth range of 1300–1800.

Models are usually updated with new obtained samples and information in machine learning. This updating process usually improves the model performance over time. As the next step of our experiment, the update process is simulated on well 202-107 to investigate its impact on the presented model performance. In this simulation, after finding the information related to every five new meters of the drilled well, the model is updated with obtained samples. The updated model will estimate the proper mud weight for the next step.

The result of this simulation is shown in Figure 5. The blue line of this plot (the bottom line) refers to general errors. In other words, the general error line indicates the MAE of the validation data set during the update process with newly obtained samples. This blue line can be a benchmark for the updated model, and the green line (the top line) is related to the MAE of the model for the new five samples in each step. As shown in Figure 5, the MAE is gradually reduced from 1.15 to 0.09. In a nutshell, in this case, the model performance significantly improves with the updating process.

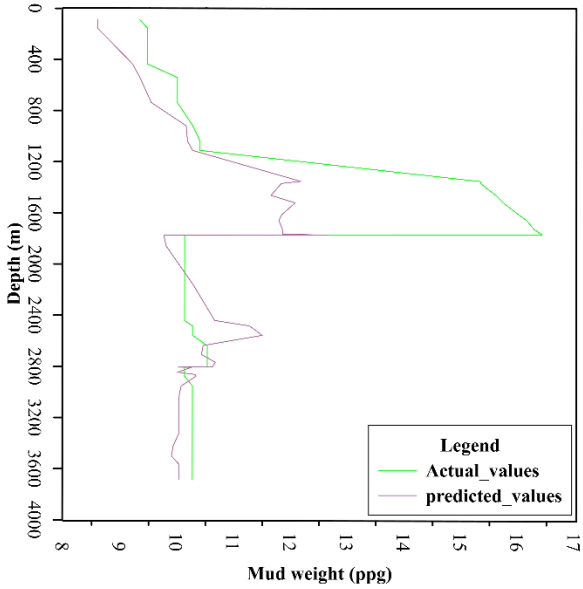

Figure 4
Comparison of actual and estimated mud weights at different depths for well 202-107.

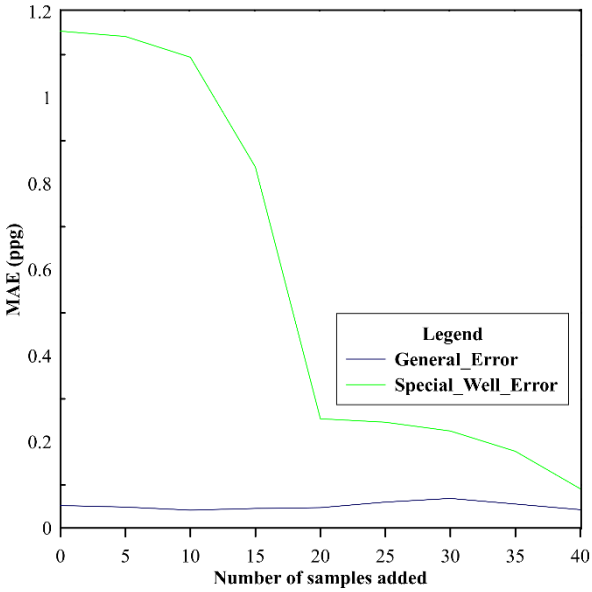

Figure 5
The trend of the model during the updating process for well 202-107 of Iran.

The total number of samples of this well is 50. The samples of the Iranian wells are too limited compared to the British and Norwegian wells, which have between 1000 and 6000 samples. Therefore, one of the sampling techniques is used as the next step to improve the obtained model.

Oversampling and undersampling in data analysis are the main categories of sampling techniques used to adjust the distribution of a data set (Domingues et al., 2018). One of the simplest oversampling methods is applied to balance the small amount of Iranian field data compared with the British and Norwegian data. This method is to repeat the Iranian samples. Therefore, the samples of the Iranian field have been repeated 100 times in our data set. It should be noted that the number of repetitions has been obtained by trial and error. The comparison of actual and estimated mud weights at different depths has been again made for well 202-107 (Figure 6). As can be seen, in this new experiment, the model prediction in a depth range of 1300 to 1800 has a significant improvement compared to Figure 4.

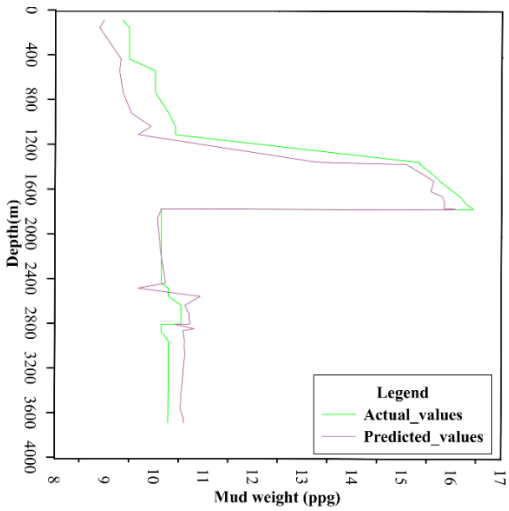

Figure 6
Comparison of actual and estimated mud weights at different depths for well 202-107 after applying the oversampling technique.

6. Conclusions

As one of the main sectors of modern industry that aims to increase production, intelligent manufacturing decreases costs, reduces breakdown time, and produces intelligent and sustainable products. This research has taken a small step toward intelligent manufacturing in the oil and gas industry.

Drilling mud with proper weight prevents the instability of the wellbore and reduces the risk of damage to the formation. Thus, this work presents a mud weight estimator for Iranian wells using deep-learning techniques. Using deep learning in the oil and gas field is a nascent approach.

The data set is integrated from three fields in the United Kingdom, Norway, and Iran. In our experiments, the MAE average of the presented mud weight estimator is 0.03849, indicating that, on average, the estimated mud weight values only have a ±0.03849 pound per gallon difference from the actual values, which is acceptable considering the operational conditions.

Our proposed model has been challenged using one well of Iran by simulating what happens in the real world in the experiments. The estimated mud weights for the unseen new wells are acceptable; significantly, if the model is updated with the samples of new drilled depths, the model will perform better for the following depths of the well. The improvement trend in the updated and oversampled models shows that the model learns many common patterns among all wells and can also learn the specific patterns of each well during the drilling process.

The proper mud weight is not a specific value in the real world, so an acceptable and safe interval (window) is defined. In other words, if the mud weight is selected from this safe window, the drilling most probably does not confront some problems, such as Kick, Gain, or Lost Circulation problems. In OGA, Volve, and PDF data sets, only the used mud weight has been recorded in most cases. Thus, the only way to check the accuracy of our model is to compare the estimated values with the recorded values. Therefore, even if the estimated mud weight is not the same as the recorded weight, it is very likely that the estimated value can be in the safe window. In other words, our presented model is expected to have better performance in actual drilling conditions.

In a nutshell, a drilling mud weight estimator is presented for Iranian wells in this research using deep-learning techniques. There are two main types of drilling mud in the oil and gas industry: water-based and oil-based. This work did not distinguish between these two types of drilling muds. As a future work, we will focus on water-based and oil-based muds separately since this separation may improve our estimators.

Nomenclature

AI	Artificial intelligence
ANN	Artificial Neural Network
ECD	Equivalent circulating density
MAE	Mean Absolute Error
ppg	Pound per gallon
RELU	Rectified Linear Unit
PDF	Pars Drilling Fluids Company
S_v	Vertical stress
P_p	Pore pressure

σ'_v	Vertical effective stress
NDR	National Data Repository
OGA	UK Oil and Gas Authority
UKCS	UK Continental Shelf

Acknowledgments

We would like to thank the PDF Company for providing the Iranian field data set and their technical advice during this research.

References

Aadnoy, B. S., and Belayneh, M., Elastoplastic Fracturing Model for Wellbore Stability Using Non-Penetrating Fluids, Journal of Petroleum Science and Engineering, Vol. 45, No. 3-4, p. 179–192, 2004.

Aadnoy, B. S., and Chenevert, M. E., Stability of Highly Inclined Boreholes (Includes Associated Papers 18596 and 18736), SPE Drilling Engineering, Vol. 2, No. 04, p. 364–374, 1987.

Abbas, A. K., Almohammed, H. H., Alqatran, G., Mohammed, H. Q., and Mohammed, A., Determination of Safe Operating Mud Weight Window from Well Logging Data Using Machine Learning Algorithms, Offshore Technology Conference Asia, Offshore Technology Conference, 2020.

Ahmed S, A., Elkatatny, S., Ali, A. Z., Abdulraheem, A., and Mahmoud, M., Artificial Neural Network ANN Approach to Predict Fracture Pressure, SPE Middle East Oil and Gas Show and Conference, Onepetro, 2019.

Aird, P., Deepwater Drilling, Gulf Professional Publishing, Elsevier, Vol. 441, p. 475, 2019.

Alsiyabi, K., AL-Aamri, M., and Siddiqui, N. A., Effective Geomechanics Approach for Wellbore Stability Analysis, SPE Middle East Oil and Gas Show and Conference, Onepetro, 2019.

Bakhtiarizadeh, S., Asadi, I. S., and Sadat, A. Y., Solving Bore Hole Instability Problem of Underbalanced Drilling in Several Wells of an Iranian Oil Field, Middle East Drilling Technology Conference, and Exhibition, Onepetro, 2009.

Bandura, L., Halpert, A. D., and Zhang, Z., Machine Learning in The Interpreter's Toolbox: Unsupervised, Supervised, and Deep-learning Applications, 2018 SEG International Exposition and Annual Meeting, Onepetro, p. 4633–4637, 2018.

Bataee, M., Irawan, S., and Kamyab, M., Artificial Neural Network Model for Prediction of Drilling Rate of Penetration and Optimization of Parameters, Journal of The Japan Petroleum Institute, Vol. 57, No. 2, p. 65–70, 2014.

Behnoud Far, P., and Hosseini, P., Estimation of Lost Circulation Amount Occurs During Under Balanced Drilling Using Drilling Data and Neural Network, Egyptian Journal of Petroleum, Vol. 26, No. 3, p. 627–634, 2017.

Cheatham, J. B., Wellbore Stability, Journal of Petroleum Technology, Vol. 36, No. 06, p. 889–896, 1984.

Chollet, F., Deep Learning with Python, Simon and Schuster, 2017.

Chollet, F., Keras, Https://Github.Com/Fchollet/Keras/, 2015.

De Felice, F., Petrillo, A., and Zomparelli, F., Prospective Design of Smart Manufacturing: an Italian Pilot Case Study, Manufacturing Letters, Vol. 15, p. 81–85, 2018.

Deng, S., Fan, H., Tian, D., Liu, Y., Zhou, Y., Wen, Z., and Ren, W., Calculation and Application of Safe Mud Density Window in Deepwater Shallow Layers, Offshore Technology Conference, One petro, 2016 .

Domingues, I., Amorim, J. P., Abreu, P. H., Duarte, H., and Santos, J., Evaluation of Oversampling Data Balancing Techniques in The Context of Ordinal Classification. 2018 International Joint Conference on Neural Networks (IJCNN), IEEE, p. 1–8, 2018.

Equinor-Company, Volve Dataset, Https://Www.Equinor.Com/, 2021.

García, S., Luengo, J., And Herrera, F., Data Preprocessing in Data Mining, Cham, Switzerland, Springer International Publishing., Vol. 72, 2015.

Han, Y., and Meng, F. F., Selecting Safe Mud Weight Window for Wellbore in Shale While Drilling Using Numerical Simulation, IADC/SPE Drilling Conference and Exhibition, Onepetro, 2014.

Hareland, G., and Dehkordi, K. K., Wellbore Stability Analysis in UBD Wells of Iranian Fields, SPE Middle East Oil and Gas Show and Conference, One petro, 2007.

Kamgue Lenwoue, A. R., Deng, J., Liu, W., Li, H., Palencia Yrausquin, E., and Matamba, G., Wellbore Stability Analysis Using an Elastoplastic Mogi–Coulomb Model, 53rd US Rock Mechanics/Geomechanics Symposium, One petro, 2019.

Khatibi, S., Farzay, O., and Aghajanpour, A., A Method to Find Optimum Mud Weight in Zones with No Safe Mud Weight Windows, 52nd US Rock Mechanics/Geomechanics Symposium, Onepetro, 2018 .

Kim, P., MATLAB Deep Learning. With Machine Learning, Neural Networks and Artificial Intelligence, Vol. 130, No. 21, 2017.

Kim, S., Kang, I., And Kwak, N., Semantic Sentence Matching with Densely-connected Recurrent and Co-Attentive Information, Proceedings of The AAAI Conference on Artificial Intelligence, Vol. 33, No. 01, P. 6586–6593, 2019.

Li, W., Classifying Geological Structure Elements from Seismic Images Using Deep Learning, 2018 SEG International Exposition and Annual Meeting, One petro, p. 4643–4648, 2018.

Li, Y., Sun, R., and Horne, R., Deep Learning for Well Data History Analysis, SPE Annual Technical Conference and Exhibition, Onepetro ,2019.

Liu, Z., Zhou, F., Feng, X., Yang, Z., and Fan, F., Calculation of Drilling Mud Density Window and Open-hole Wellbore Extension Limit for Extended Reach Wells, 52nd US Rock Mechanics/Geomechanics Symposium, Onepetro, 2018.

Mclellan, P. J., Assessing the Risk of Wellbore Instability in Horizontal and Inclined Wells, Journal of Canadian Petroleum Technology, Vol. 35, No. 05, 1996.

Meyers, R. A., Handbook of Petroleum Refining Processes, McGraw-Hill Education, 2016 .

NDR, UK National Data Repository, Https://Ndr.Ogauthority.Co.Uk/, 2021.

Norwegian-petroleum, Volve Field, Https://Www.Norskpetroleum.No/En/Facts/Field/Volve/, 2021.

Oil And Gas Authority (OGA), National Data Repository (NDR), Https://Www.Ogauthority.Co.Uk/, 2021.

Park, R. G., and Park, G., Foundations of Structural Geology, Glasgow and London: Blackie, p. 148, 1989.

Pars Drilling Fluids (PDF), Http://Www.Parsdrill.Com/, 2021.

Pattillo, P. D., Elements of Oil and Gas Well Tubular Design, Gulf Professional Publishing, 2018 .

Peng, S., and Zhang, J., Engineering Geology for Underground Rocks, Springer Science and Business Media, 2007.

Pereira, L. G., Gandelman, R. A., Clemente, R. G., Teixeira, P. H. S., And Teixeira, G. T., Development of Software to Predict Mud Weight for Pre-salt Drilling Zones Using Machine Learning, OTC Brazil, One petro, 2013.

Phan, D.T., Liu, C., Altammar, M.J., Han, Y., And Abousleiman, Y.N., Application of Artificial Intelligence to Predict Time-dependent Safe Mud Weight Windows for Inclined Wellbores, International Petroleum Technology Conference, One petro, 2020.

Raschka, S., and Mirjalili, V., Python Machine Learning: Machine Learning and Deep Learning with Python, Scikit-learn and TensorFlow, Second Edition, 2017.

Sebtosheikh, M. A., Motafakkerfard, R., Riahi, M. A., and Moradi, S., Separating Well Log Data to Train Support Vector Machines for Lithology Prediction in A Heterogeneous Carbonate Reservoir. Iranian Journal of Oil and Gas Science and Technology, Vol. 4, No. 2, p. 1–14, 2015.

Shadizadeh, S. R., Karimi, F., and Zoveidavianpoor, M., Drilling Stuck Pipe Prediction in Iranian Oil Fields: An Artificial Neural Network Approach, Iran J. Chem. Eng., Vol. 7, No. 4, P. 29–41, 2010.

Shen, X., Williams, K., and Standifird, W., Enhanced 1D Prediction of The Mud Weight Window for A Subsalt Well, Deepwater GOM, SPE Deepwater Drilling and Completions Conference, Onepetro, 2014.

Silva, C., Rabe, C., And Fontoura, S., Geomechanical Model and Wellbore Stability Analysis of Brazil's Pre-Salt Carbonates, A Case Study in Block BMS-8, OTC Brazil, Onepetro, 2017.

Tewari, S., Assessment of Data-driven Ensemble Methods for Conserving Wellbore Stability in Deviated Wells, SPE Annual Technical Conference, and Exhibition, Onepetro, 2019.

Tuptuk, N., and Hailes, S., Security of Smart Manufacturing Systems, Journal of Manufacturing Systems, Vol. 47, p. 93–106, 2018.

Yavari, H., Sabah, M., Khosravanian, R., and Wood, D., Application of an Adaptive Neuro-fuzzy Inference System and Mathematical Rate of Penetration Models to Predicting Drilling Rate, Iranian Journal of Oil and Gas Science and Technology, Vol. 7, No. 3, p. 73–100, 2018.

Zahiri, J., Abdideh, M., And Ghaleh Golab, E., Determination of Safe Mud Weight Window Based on Well Logging Data Using Artificial Intelligence, Geosystem Engineering, Vol. 22, No. 4, p. 193–205, 2019.

Zeynali, M. E., Mechanical and Physico-chemical Aspects of Wellbore Stability During Drilling Operations, Journal of Petroleum Science and Engineering, Vol. 82, p. 120–124, 2012.

Zhou, H., Niu, X., Fan, H., And Wang, G., Effective Calculation Model of Drilling Fluids Density and ESD For HTHP Well While Drilling, IADC/SPE Asia Pacific Drilling Technology Conference, Onepetro, 2016.

Zoback, M. D., Barton, C. A., Brudy, M., Castillo, D. A., Finkbeiner, T., Grollimund, B. R., Moos, D. B., Peska, P., Ward, D. J., and Wiprut, D. J., Determination of Stress Orientation and Magnitude in Deep Wells, International Journal of Rock Mechanics and Mining Sciences, Vol. 40, No. 7–8, P. 1049–1076, 2003.