Italian Journal of Engineering Geology and Environment
www.ijege.uniroma1.it

ESTIMATION BRITTLENESS INDEX IN CARBONATE ENVIRONMENTS USING LOG AND LITHOLOGY DATA AND DEEP LEARNING TECHNIQUES

Farhad MOLLAEI(*), Ali MORADZADEH(*) & Reza MOHEBIAN(*)

(*)*School of Mining, College of Engineering, University of Tehran - Tehran (Iran)*
Corresponding author: mohebian@ut.ac.ir

EXTENDED ABSTRACT

Uno dei parametri geomeccanici più importanti nella meccanica delle rocce, è la fragilità che ha un grande impatto sui processi di rottura e sulle attività di perforazione. In generale, la fragilità è funzione della resistenza della roccia alla deformazione in campo elastico. I metodi per valutare la fragilità delle rocce sono principalmente divisi in tre categorie: (1) basati su moduli elastici, (2) misurazioni dirette di sforzo-deformazione in laboratorio e (3) contenuto di minerali. L'indice di fragilità dinamica BI_{dyn} è ottenuto dai parametri del modulo di Young dinamico e del rapporto di Poisson dinamico. L'indice di fragilità statica BI_{st} può essere ottenuto dalla resistenza alla compressione uniassiale e dalla resistenza alla trazione calcolate in laboratorio. La resistenza alla trazione è generalmente considerata pari a 0.1 della resistenza alla compressione uniassiale. Inoltre, è possibile determinare l'indice di fragilità dei minerali BI_{lit} e in questo campo sono state proposte varie relazioni empiriche.

In questo articolo, erano disponibili i dati dei log SGR, CGR, DT, RHOB, RT, NPHI, CALIPER, V_p, PEF per determinare il *BI* dinamico e statico utilizzando algoritmi di *deep learning*. Per selezionare le caratteristiche effettive e gli input adeguati agli algoritmi, il coefficiente di correlazione delle caratteristiche dovrebbe essere controllato con il *BI*. Uno dei metodi per selezionare i parametri è l *Auto-encoder deep network*, in cui i registri V_p, RHOB, CALIPER, PEF e CGR sono stati selezionati come input degli algoritmi.

In questo studio erano disponibili 16330 dati nell'intervallo di profondità compreso tra 3551 e 3799 metri. Inizialmente, il 24% di questi dati sono stati separati dalla fine del *dataset* come dati ciechi e dgli altri 12330 dati, l'80% è stato selezionato come dati di addestramento e il 20% come dati di test. E per evitare un adattamento eccessivo, è stata considerata una suddivisione di convalida dello 0.1 dei dati di addestramento. Nella fase successiva, è stata eseguita la normalizzazione dei dati per ottenere una maggiore precisione attraverso la funzione di normalizzazione Min-Max, che regola i dati disponibili tra zero e uno. Di seguito, per l'ottimizzazione è stata utilizzata la funzione Adam Optimizer. Per valutare il modello e confrontare i risultati degli algoritmi di *deep learning*, sono stati utilizzati MAPE, MAE, MSE, RMSE, NMSRE e R^2.

Per stimare BI_{dyn}, i valori di errore per i dati ciechi sono uguali a MSE_{MLP}=5.1687, MSE_{LSTM}=1.7135, MSE_{CNN}=1.0292 e i valori R^2 sono uguali a R^2_{MLP}=0.9199, R_{LSTM} =0,9734, R^2_{CNN}=0,9841. Secondo i risultati dell'R^2, per i dati ciechi in tre algoritmi, si può concludere che il metodo più robusto e migliore di previsione del BI_{dyn} è l'algoritmo CNN mentre il più debole è l'algoritmo MLP. Continuando, il BI_{dyn} viene convertito in BI_{st}, con errori e accuratezza basati sui dati ciechi. Per stimare il BI_{st}, i valori di errore per i dati ciechi sono uguali a MSE_{MLP}=1.4823, MSE_{LSTM}=0.9425, MSE_{CNN}=0.8444 e i valori R^2 sono uguali a R^2_{MLP}=0.9457, R^2_{LSTM}=0,9655, R^2_{CNN}=0,9691. In base ai risultati e considerando gli errori di previsione e l'R^2, per i dati ciechi in tre algoritmi, si può concludere che il metodo più robusto e migliore per il $BI_{dyn,}$ poi convertito per la predizione del $BI_{st,}$ è l'algoritmo CNN e il più debole in questo caso è legato all'algoritmo MLP.

Per stimare BI_{st}, i valori di errore per i dati ciechi sono uguali a MSE_{MLP}=1.2071, MSE_{LSTM}=0.9076, MSE_{CNN}=0.6719 e i valori dei coefficienti di determinazione sono uguali a R^2_{MLP}=0.9557, R^2_{LSTM}=0,9667, R^2_{CNN}=0,9754. In base ai risultati e considerando gli errori di previsione e l'R^2, per i dati ciechi in tre algoritmi, si può concludere che il metodo più robusto e migliore di previsione del BI_{st} è l'algoritmo CNN e le prestazioni più deboli in questo caso correlate, invece, all'algoritmo MLP.

Di seguito sono stati confrontati tutti i BI_{dyn} misurati e tutti i BI_{dyn} previsti. Inoltre, sono stati confrontati tutti i dati BI_{st} previsti e i dati fondamentali e ampliato il confronto tra campioni fondamentali e dati BI previsti e poi con il BI ottenuto dalla percentuale in volume di minerali. I risultati mostrano che le varietà di arenaria hanno un BI più elevato rispetto ai carbonati. Secondo i risultati, c'è un'ottima corrispondenza tra i dati principali e il BI_{st} previsto. Anche secondo il confronto tra il BI ricavato dalla percentuale in volume di minerali e i BI_{dyn} e BI_{st} ottenuti, c'è una buona corrispondenza. Inoltre, secondo il BI calcolato dalla litologia, si può affermare che in ambiente carbonatico, nelle aree in cui la litologia è calcarea e dolomitica, si ha un BI inferiore rispetto alle aree in cui la litologia è arenaria.

Il confronto dei risultati ottenuti da questi tre algoritmi, mostra che gli algoritmi LSTM e CNN hanno prestazioni adeguate per prevedere il BI, che può essere utilizzato semplicemente disponendo di log regolari, perché l'errore ottenuto e l'R^2 hanno valori accettabili sia per i dati di test che ciechi. Quindi, in generale, si può dire che gli algoritmi di *deep learning* sono un metodo efficiente per prevedere il BI_{dyn} e il BI_{st}.

IJEGE, 1 (2024): 49-66, DOI: 10.4408/IJEGE.2024-01.O-04
E-ISSN 2035-5688 | ISSN 1825-6635 / ©Author(s)

ABSTRACT

Brittleness index is one of the most important Geomechanical parameters, which has a great impact on the rock breaking process and drilling activities. The methods of evaluating brittleness of rocks are mainly divided into three categories: (1) direct laboratory, (2) mineral content, and (3) based on elastic moduli. one of the efficient methods for brittleness index predicting is use of intelligent methods, which are low-cost and accurate methods, and it is possible to predict the brittleness index using log and lithology data. In this study, dynamic and static brittleness index values are predicted using deep learning (DL) algorithms and lithology data in carbonate environment in one of the hydrocarbon fields in southern part of Iran. In this paper, the effective features were selected using the deep learning algorithm of the Auto-encoder, and the dynamic and static brittleness index was estimated using the MLP, LSM, and CNN algorithms. As 12 laboratory core samples were available, at first the brittleness index were calculated by relevant empirical relations and data of some available well logs in order to generalize these core results to the entire target depth range of 3551.07 to 3799.78 meters. Then a set of relationship between the well's logs derived dynamic and static brittleness index and laboratory results was determined for the depths where the laboratory samples were recorded. Following that, an Auto-encoders deep network was used to select the effective features in predicting the brittleness index, and finally by using MLP, LSTM and CNN networks the value of dynamic and static brittleness index was predicted. Here, the goal is to obtain the brittleness index values with high accuracy wherein there no core data. The performance of the three algorithms prediction models is tested by blind data sets that the models have not seen before. Furthermore, the results were checked and evaluated by set of statistical measures like MAE, MAPE, MSE, RMSE, NRMSE and R^2 values that calculated for train, test and blind dataset. At first, dynamic brittleness index estimate using log data and three algorithms and R^2 for blind data equal to R^2_{MLP}=0.91, R^2_{LSTM}=0.97, R^2_{CNN}=0.98, in the following, using MLP, LSTM and CNN the dynamic brittleness index has been converted into a static brittleness index and R^2 for blind data equal to R^2_{MLP}=0.94, R^2_{LSTM}=0.96, R^2_{CNN}=0.96. Finally, the static brittleness index has been estimated directly from the log data without the relation of dynamic to static transformation and R^2 for blind data equal to R^2_{MLP}=0.95, R^2_{LSTM}=0.96, R^2_{CNN}=0.97. Finally, the dynamic and static brittleness index was compared with the brittleness index obtained from lithology, and there is a good match between them. The results show that the deep learning algorithm is a novel method, robustness and accurate method in estimating the dynamic and static brittleness index using conventional logs. The results show used CNN and LSTM networks as new deep learning algorithms to predict brittleness index.

***Keywords**: dynamic and static brittleness index, log data, deep learning algorithm, geomechanical parameters*

INTRODUCTION

One of the important Geomechanical parameters is brittleness and it is one of the most important characteristics of rock mechanics, which has a great impact on the rock breaking process and drilling activities. In general, the rock with more brittleness properties shows lower plastic deformation values at the time of failure compared to softer rocks (Morley, 1954). In other words, in rock mechanics, brittleness refers to a type of fracture during which there is no permanent deformation (Ramzi, 1976). Morley (1954) expressed brittleness as the absence of ductility (elastic state under the influence of stress). Hetni (1966) defined brittleness as lack of ductility (opposite of ductility). Ramzi (1967) expressed one of the most important definitions for the brittleness index based on adhesion. In this definition, when the internal adhesion of rock materials that deform in the elastic range decreases or disappears, it is said that the rock has a brittle behavior. Obert and Duval (1967) called materials such as cast iron and many rocks, which usually break in tension or only shortly after tension, as brittle materials. In other words, brittleness can be defined as one of the properties of materials that break tensile with some flow or without using flow. Huka and Das (1974) stated that the concept of Brittleness is not precisely stated, but with these interpretations, what was obtained for rocks with a high Brittleness value is: low elasticity, tensile failure, high ratio Compressive strength to tensile strength, high elasticity, high internal friction angle. Altindag (2003) defined the tendency of rock to break along with the lack of ductility that can be seen at low stresses as brittleness. In general, the property of brittleness has a significant effect on the process of breaking rocks. Xie *et alii* (2008) suggested that the deformation and failure behavior of rock is the destabilizing phenomenon accompanied by the storage, dissipation, transformation, and release of energy. Brittleness index and compressive strength of rock are effective parameters in the operation of the device during drilling. Rocks exhibit brittle behavior under compression and tension, therefore determining rock brittleness is of great importance in practical projects of rock mechanics, underground structures and nuclear waste tanks (Yagiz, 2009). In general, brittleness is a function of strength and indicates the resistance of rock against deformation in the elastic range (Altindag, 2010). However, there is no direct and standard method for measuring brittleness, but it can be indirectly used to determine the concept of fragility by using some relations about rock, such as different ratios of compressive and tensile strength of rock (Guktan, 1991; Lawn & Marshall, 1979). Consequently, the brittleness index is largely a function of rock properties and freeze-thaw cycles in cold regions; therefore, its measurement based on these factors is highly crucial. Rock brittleness is closely influenced by the mineral composition (Zhang *et alii*, 2017), physical properties (Qian *et alii*, 2017), loading condition (Xiao *et alii*, 2020), micro-fracture (Wang *et*

alii. 2021), and mechanical characteristics (Li *et alii*, 2020; Xu *et alii*, 2021). The methods of evaluating brittleness of rocks are mainly divided into three categories: (1) based on elastic moduli, (2) direct laboratory stress-strain measurements, and (3) mineral content. The dynamic brittleness index (BI_{dyn}) is obtained from the parameters of dynamic Young's modulus (E_{dyn}) and dynamic Poisson's ratio (υ_{dyn}), which are shown in relation 3, respectively dynamic brittleness index (Valin & Antia, 2020).

$$BI_{dyn} = E_{dyn} + \frac{\upsilon_{dyn}}{2} \quad (1)$$

The static brittleness index (BI_{st}) can be obtained from the uniaxial compressive strength (σ_c) and tensile strength (σ_t) that are calculated in the laboratory. Tensile strength is usually considered 0.1 uniaxial compressive strength. Equation 2 shows the static brittleness index (Altindag, 2010).

$$BI_{st} = \sqrt{\frac{\sigma_c \times \sigma_t}{2}} \quad (2)$$

Also, the brittleness index of minerals (BI_{lit}) can be determined, and in this field, various relations have been presented, one of which is shown in relationship 3, which is related to the parameters of quartz volume (V_{sand}), clay volume (V_{cl}), and carbonate volume (V_{ca}) (Jarvie, 2007).

$$BI_{lit} = \frac{V_{sand}}{V_{sand}+V_{cl}+V_{ca}} \quad (3)$$

Today, machine learning and deep learning are widely used in earth sciences such as seismology (Mohebian *et ali*, 2017), petrophysics (Mohebian *et alii*, 2019, 2021) and geothermal (Kelishami *et alii*, 2022). In recent years, rock brittleness has been extensively studied by many researchers in the geo-mechanics filed. However, the definitions of brittleness are still ambiguous, an agreement of the measurement standards of brittleness has not yet been reached (Kahramana, 2002; Altindag, 2003; Goktan & Yilmaz, 2005; Wang & Gale, 2009; Wang *et alii*, 2015; Zhang *et alii*, 2016). Jarvie *et alii* (2007) propose a brittleness equation based on the amount of quartz, calcite, and clay minerals, in which quartz is considered to be the brittle mineral. Rickman *et alii*. (2008) propose an average brittleness equation based on the elastic parameters of Poisson's ratio and Young's modulus. Perez (2013) compare the brittleness index (*BI*) estimated from mineral content and brittleness average estimated from elastic parameters. Da Silva (2013) finds that the *BI* computed from mineral content is positively correlated to $\mu\rho$ and negatively correlated to $\lambda\rho$. Jin *et alii* (2014) review brittleness estimation from geomechanical and petrophysics consideration. Zhang *et ali*i (2014b) find that inverted results from conditioned gathers have better resolution and higher correlation coefficients with well logs. Shi *et alii* (2016) estimated the brittleness index using machine learning extreme and back propagation artificial neural network. Ghobadi and Naseri (2016) predicted the brittleness index using artificial neural network and Multiple Regression Models. Quian *et alii* (2019) investigated the brittleness index of shale using the average weight-day hill. Ore and Gao (2021) predicted the brittleness index using artificial neural network and support vector regression. Zhang *et alii* (2022) doing Rock Brittleness Evaluation Index based on Ultimate Elastic Strain. Energy. In the research of Ghobadi *et alii* (2023) relationship between rock brittleness is investigated by using the ratio of point index to porosity (PMP) in the sandstones of the Qom Formation. In addition, the available estimation methods for the rock brittleness index are summarized and their application is briefly discussed. Wang *et alii* (2023) a digital drilling method are proposed to evaluate the mechanical anisotropy of rock and the anisotropic effect on cutter wear. In considering critical friction, the cutting efficiency and contact stress were determined from the revised drilling model in order to characterize the drilling process. Wang *et alii* (2023) analyzes the three types of brittleness indices in detail, A stress–strain curve-based brittleness index would be appropriate for the evaluation of brittleness if the parameters represented unique and certain stress–strain curves. Therefore, a new brittleness index is proposed. It consists of two components: the stress variation index and the strain variation index, which reflects the strain ratio of the stress drop and the relationship between elastic and post-peak strains, combining the stress variation and strain variation in the pre- and post-peak phases. In the research of Rahimi Shahid *et alii* (2023) the uniaxial compressive strength values of zone 38 (western Iran) were extracted and analyzed. Then, dry and saturated brittleness index were determined using existing experimental relationships and dry and saturated uniaxial compressive strength. Wang *et alii* (2023) proposed a new method for evaluating rock brittleness using a digital drilling approach. A cutting model was established to describe the relationship between the energy characteristics and mechanical parameters of the rock during the drilling process, accounting for the effects of friction and drilling fluid. Zhang *et alii* (2024) used the reflection coefficient equation for the direct inversion of the brittleness index considering the cumulative error in the indirect inversion. This article introduces a time-frequency mixed domain inversion (TFDI) method, considering that the frequency domain inversion (FDI) method has excellent resolution and the time domain inversion (TDI) method has good noise immunity. Asemi *et alii* (2024) introduced new index for describing rock brittleness using crack initiation and crack damage stress thresholds for rocks with Class I stress-strain curves. Uniaxial compressive strength (UCS) tests were conducted on fine-grained, medium-grained, and coarse-grained dolomite rock specimens in order to evaluate the performance of this index on describing brittleness. Fang *et alii* (2024)

introduced a multi-mineral composition equivalent model for complex lithologies that enables the accurate calculation of V_p and V_s These ratios serve as the foundation for pre-stack elastic parameter predictions, which include Poisson's ratio and Young's modulus. By comparing the predicted parameters with well-logging measurements.

Considering the high importance of brittleness index in geomechanical modeling and hydraulic fracturing, a method that can predict brittleness index with high accuracy is very important. One of the methods used are intelligence methods. According to previous researches, many methods have been performed to estimate geomechanical parameters and brittleness index. But because accurate estimation of these parameters is very important for modeling and geomechanical studies, this study try to introduce a new algorithm for estimating this parameter using deep learning algorithm and well log data with high accuracy and low cost. On this basis, three algorithms are used to predict brittleness index (multilayer perceptron network (MLP), transformed convolutional neural network (CNN) and one kind of recurrent neural network named as LSTM). It aims to establish a model that can be used to estimate dynamic and static brittleness index according to conventional petrophysical well logs. In this paper, dynamic and static brittleness index values are predicted using deep learning (DL) algorithms and lithology data in carbonate environment in one of the hydrocarbon fields in southern part of Iran. The effective features were selected using the deep learning algorithm of the Auto-encoder, and the dynamic and static brittleness index was estimated using the MLP, LSM, and CNN algorithms. In the first step, the Auto-encoders deep network was used to select the effective features. Subsequently, the selected logs were utilized as input parameters of the model to predict the brittleness index using three MLP, LSTM and CNN algorithms. Finally, the results of the three algorithms were compared with each other. In three separate steps, the brittleness index has been estimated using the deep learning algorithm. First, the dynamic brittleness index is estimated with deep learning algorithm. In the following, the dynamic brittleness index has been converted into a static brittleness index with deep learning algorithms. In the end step, the static brittleness index is estimated using deep learning algorithms directly with conventional logs and finally, the dynamic and static brittleness index was compared with the brittleness index obtained from lithology. The obtained results show the effective use of deep learning algorithms (CNN and LSTM), in predicting brittleness index.

METHODS

Multi-layer perceptron neural network

The multi-layer perceptron neural network is one of the feedforward artificial neural networks. In a multilayer perceptron neural network, there will be at least three layers of nodes: input layer, hidden layer, output layer. Neural network nodes are the computing units in a neural network. In this neural network, the outputs of the first layer (input) are used as the inputs of the next (hidden) layer; This work continues in this way, until, after a certain number of layers, the outputs of the last hidden layer are used as the inputs of the output layer. The input and output layer are called the hidden layer (Alavi *et alii,* 2010). Multi-layer perceptron networks also contain a set of weights that must be adjusted for neural network training and learning.

Recurrent neural network

The recurrent neural network is a neural network in which the output of each cell is not only related to the previous layers, the output of each cell also feeds its own input; In other words, the recursive network by using a recursive loop makes it possible for the information obtained (last state) by the network at the end of each calculation to be used for future calculations. One of the problems of recurrent neural networks is their inability to learn long-term dependence. To solve this problem, Hochreiter *et alii* (1997) proposed an architecture for recurrent neural networks, which became known as the long-short-term memory (LSTM) architecture, and was able to compensate for the mentioned shortcoming (Hochreiter & Schmidhuber, 1997). Since the introduction of LSTM, improvements have been made to increase its efficiency. Figure 1 shows the structure of an LSTM block, which includes three inputs, output and forget gates. The output of the LSTM block is recursively connected to the input of the block and the input of the three mentioned gates. Activation functions in input and output gates are usually hyperbolic tangent function and activation function in forgetting gate is sigmoid function.

Convolutional neural network

Two This method was introduced in the 1980s and 1990s. Although the convolutional neural network was forgotten for a while, however, since 2012 and with the modification of the convolutional neural network (CNN or ConvNet), it has made great progress in the machine field and is expanding rapidly

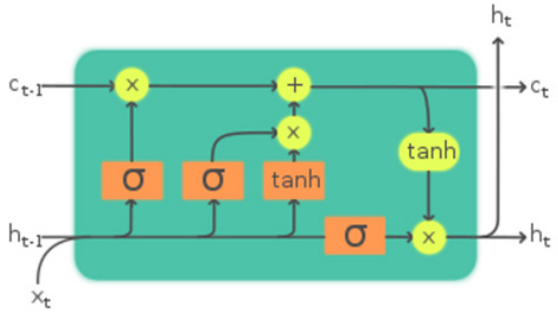

Fig. 1 - Structure of an LSTM block (Greff et alii, 2017)

(Guo *et alii*, 2015). One of the advantages of the CNN method is the reduction of the number of weight parameters with the help of the concept of weight sharing. Also, in the generation of features in this method, the neighborhoods of a pixel are considered, which is compatible with the assumption that the statistical properties of the pixels of a part in the images are the same (Guo *et alii*, 2015). In multi-layer perceptron neural networks, the neurons are fully connected to each other, and with the increase in the dimensions of the images, the number of connections and free parameters will increase greatly, and as a result, large training data will be required. Therefore, during training, the problem of overfitting may occur due to the inappropriate number of training samples (Lecun *et alii*, 1998). Also, perceptron multilayer neural networks do not show resistance against noise. One of the most important weaknesses of MLP neural networks is not providing a mechanism for automatic feature generation. According to MLP problems, CNN methods have been presented. CNN consists of three layers: convolution, pooling and full connection. The convolution layer is created from the convolution of the input image with the help of local receiving areas, and in this layer, the concept of being hierarchical and approaching the object is realized. The movement of the filter or filters on the input bands is in the form of Toggle Movement (Gu *et alii*, 2018). Determining the number of filter banks and convolution layers and adjusting its parameters is one of the existing challenges. In the pooling layer, the image size is reduced. The main reason for applying this operator is to reduce the volume of calculations and Invariant of the features with respect to the image size (Yoo, 2015). In addition to the above advantage, the non-linearity of the features produced in the convolution stage is done with the help of the activation function in the pooling layer. The pooling operation reduces the size along the spatial dimensions, the result of which will be a mass with a smaller size. Basically, it is through this operation that a score vector is created at the end of the convolution network. Because this operation is applied to all slices, the output dimension is the same as the input dimension to the pooling layer and does not change. There are various methods for applying pooling, but the methods of averaging and choosing the maximum with the reduction of the ratio of two and the Relu activation function are among the most common methods. After the pooling stage, the two-dimensional features created in the fully connection stage are converted into a one-dimensional vector. This layer, like the MLP neural network, provides the possibility of implementing feed-forward training for CNN (Benigo *et alii*, 2013). There are different types of activation functions. Some examples of the most famous activation functions are the sigmoid function, the hyperbolic tangent function, the Relu linear rectifier function, the exponential linear function, and the softmax function. Figure 2 shows an example of image processing by convolutional neural network.

Computing environment

All training and evaluation stages were completed using Python version 3.9.17 (https://www.python.org) with the TensorFlow (version 2.6) (Abadi *et alii*, 2016), tensorflow-keras backend (version 2.6) (Chollet, 2015), Scikit-learn (version 1.3.0) (Pedregosa *et alii*, 2011), NumPy (version 1.23) (Oliphant, 2006) and Matplotlib (version 3.7.1) (Hunter, 2007) libraries. The deep learning frameworks were implemented using the CPU of a Quadro 2000M model with a 2.4 GHz Intel Core i7 x64-based processor and 16 GB of RAM, running the 64-bit operating system.

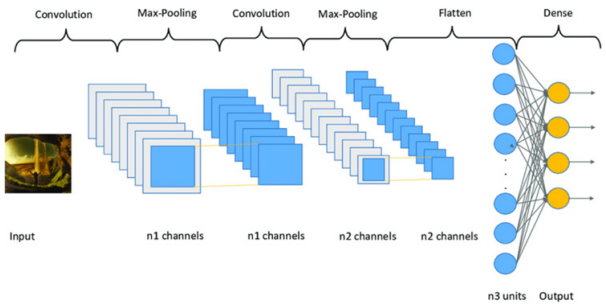

Fig. 2 - Schematic of CNN algorithm (Benigo et alii, 2013)

DATA PREPARATION

The data used is related to the well of one of the hydrocarbon fields in the southwest of Iran. This field is one of the largest oil fields in the Zagros basin, which is located in the eastern part of the structural area of Dezful embayment (Dezful embayment is a part of Zagros where most of Iran's oil fields are located. This embayment is a part of the folded Zagros, where the Asmari formation and older layers do not have a surface outcrop. In the formation of this embayment, the combined action of Qatar-Kazron right-handed fault and Balaroud left-handed fault played a major role). This field is extended with a northwest-southeast trend in the western to central part and a northeast-southwest trend in the eastern part. The surface outcrop of this field is the Aghajari formation. the Asmari formation, the Bangestan and Khami groups are the hydrocarbon reservoirs in this field. Asmari Formation is the most important reservoir rock of this field, which is divided into 6 reservoir layers. Reservoir layers one, two, three are mainly composed of dolomitic carbonates, so the density of fractures (especially in layer one) (90% dolomite) is higher. In the fourth, fifth and sixth reservoir layers of this field, due to the increase of shale and marl layers, as well as the decrease in fragility, the density of fractures decreases. The total available data are 16330, which are located in the depth range of

3551.07 to 3799.78 meters and 12 laboratory samples have been available. Figure 3 shows a schematic of the oil field. Available well logs are sum gamma ray (SGR), clean gamma ray (CGR), sonic transient time (DT), density (RHOB), resistivity (LL7) neutron porosity (NPHI), CALIPER, primary (p) and secondary (shear) velocity (V_p, V_s) and photo electric (PEF). Figure 4 shows a logs and Table 1 shows core samples. According to Figure 4 and the lithology column, the lithology is mainly carbonate with anhydrite, sandstone and shale lithologies.

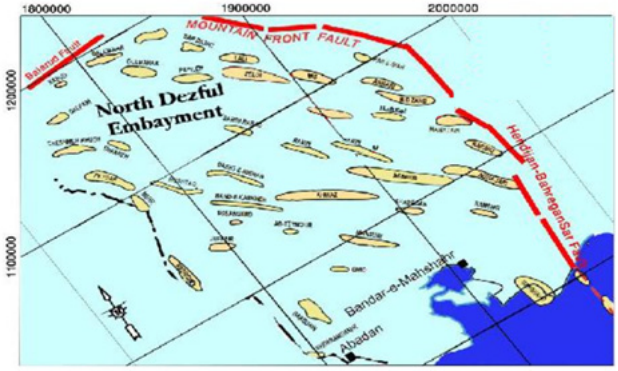

Fig. 3 - Scheme of the oil field

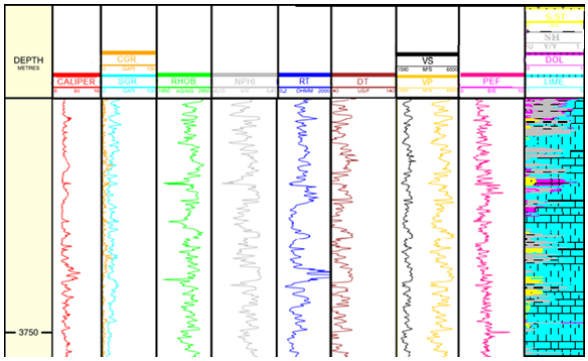

Fig. 4 - Logs in the well. from left to right: Depth, Caliper, SGR, CGR, RHOB, NPHI, RT, DT, Vp, Vs, PEF, Lithology

Number	Depth (m)	UCS (laboratory) (Mpa)
1	3553.054	90.052
2	3553.511	90.733
3	3555.644	80.518
4	3556.559	82.561
5	3557.626	75.751
6	3558.083	74.616
7	3558.997	89.825
8	3559.302	92.095
9	3560.216	92.549
10	3560.826	90.96
11	3561.588	85.512
12	3562.655	88.236

Tab. 1 - Values of laboratory samples of UCS

Feature selection to reduce dimension

The relations between brittleness index with conventional logs is a very complex nonlinear system problem, which is the result of the interaction of many elements in the earth system, that makes it difficult for us to analyze and predict brittleness index. Choosing to use the best sensitive logs instead of all available conventional logs for model training and prediction can reduce the amount of data processing and improve the processing speed and efficiency of the model (Anemangely *et alii*, 2018). Moreover, feature selection also increases the prediction precision and universal applicability of the model. In other words, if all the characteristic factors are taken into account, there will be a lot of variables input into the model, which will not only complicate the structure of the network, but also reduce the precision of the estimation results. Therefore, in order to simplify the model structure, improve the modeling ability, enhance the model prediction efficiency, and alleviate the interference of the non-main parameter variables of the model to the prediction results it is necessary to select the feature. In this paper, by using Auto-encoder deep learning algorithm, the effective features and effective logs were determined to predict brittleness index. The data of SGR, CGR, DT, RHOB, RT, NPHI, CALIPER, V_p, V_s and PEF logs were determined by applying Auto-encoder's deep learning algorithm to the main features, and logs, V_p, RHOB, CALIPER, PEF and CGR were determined as inputs for MLP, LSTM and CNN models for brittleness index, because choosing other logs would increase the error and decrease the accuracy of the model. Also, because the shear wave velocity log is not recorded in most wells, this log is not used as an input parameter of the algorithms, and the parameters of brittleness index is predicted with common logs that are recorded in all wells. Figure 5 shows the workflow of feature selection using deep learning Auto-encoder algorithm.

Dataset splits

Prediction reliability is one of the main concerns in the performance evaluation of supervised deep learning algorithms (Consonni *et alii*, 2010; Alsina *et alii*, 2017). In this study, 16330 data were available in the depth range of 3551.07 to 3799.78 meters. At first, 24% of these data were separated from the end of dataset as blind data and from the other 12330 data, 80% were selected as training data and 20% as test data. and to avoid overfitting, validation split 0.1 of the training data is considered.

Data normalization

For reduce the prediction error result from the difference of order of magnitude between the input data, the original data need to be preprocessed in the experiment. In this paper, the Min-Max Normalization method is used to normalize

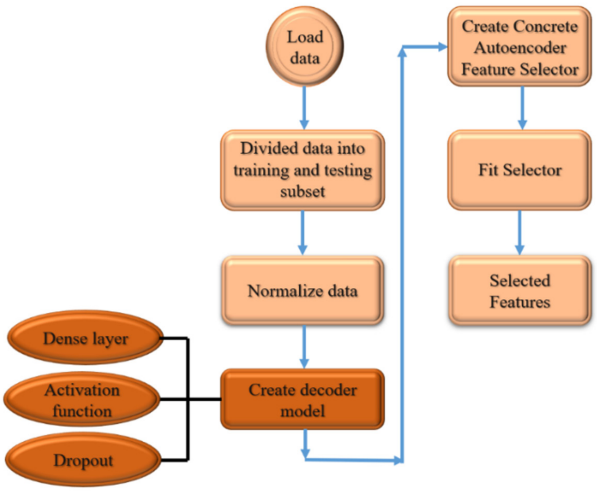

Fig. 5 - Display the workflow of feature selection using Auto-encoder technique

the original data to [0,1] range, in order to eliminate the dimensional difference.

Adam algorithm

Adam optimization algorithm is an adaptive learning rate optimization algorithm based on gradient stochastic objective function optimization (King & Ba, 2014). Adam combines the advantages of two popular optimization methods: AdaGrad (Duchi *et alii*, 2011) for sparse gradient problems and RMSProp (Tieleman & Hinton, 2012) for nonlinear and non-fixed optimization problems. The parameter update size of Adam algorithm is constant for gradient rescaling. It not only doesn't need fixed optimization objectives and is suitable for sparse gradient, but also performs step-size backtracking naturally. It has the advantages of simple implementation, with the memory required is small, high computational efficiency, etc., and is very suitable for the problems of more optimization parameters and large amount of data. Therefore, this study chooses Adam optimization algorithm to optimize deep learning algorithms.

Dynamic brittleness index using deep learning method

In the introduction, the relation of BI_{dyn} is introduced (equation 3). To estimate elastic moduli, shear wave velocity is needed, and since shear wave velocity logs are not recorded in most wells, estimating the BI_{dyn} from conventional logs is very important. BI_{dyn} was estimated with the conventional data logs of V_p, RHOB, CALIPER, PEF and CGR (the effective features were determined using Auto-encoder's deep learning algorithm) and using MLP, LSTM and CNN algorithms The log data for each case was divided into three parts, training, test and blind, according to the procedure described in dataset splits. Figure 6 shows the BI_{dyn} estimation flowchart.

The key parameters for building the MLP network in this experiment were set as follows: Batch size = 50, Learning rate = 0.01, Number of iterations = 400, Hidden layers = 2 layers, Nodes in the first hidden layer = 500, Nodes in the second hidden layer = 100, Dense layer = 1, Activation function = Relu, Optimization algorithm = Adam.

For building the LSTM network: Batch size = 256, Learning rate = 0.001, Number of iterations = 400, Hidden layers = 2 layers, Nodes in the first hidden layer = 200, Nodes in the second hidden layer = 100, Dense layer = 1, Optimization algorithm = Adam.

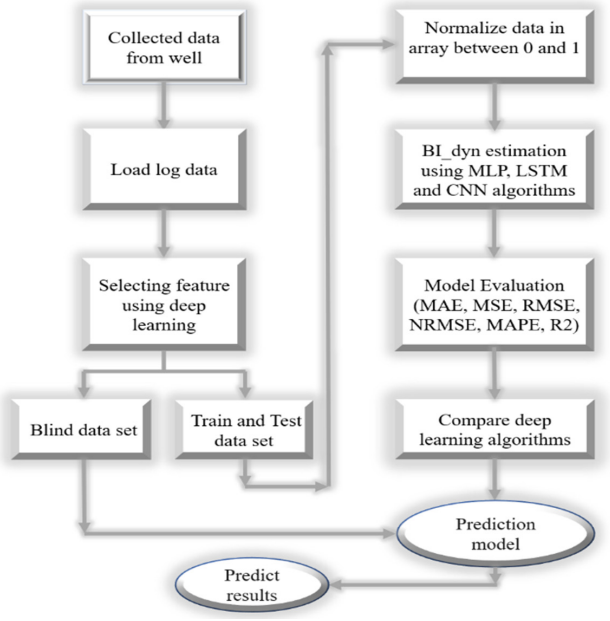

Fig. 6 - Workflow schematic for BI_{dyn} prediction using DL algorithms

Several hyperparameters were set for developing the CNN model in this study to predict the dynamic brittleness index:

Batch size = 256, Learning rate = 0.0001, Number of iterations = 100, Layers = 2 layers, Number of filters in the first layer = 128, Number of filters in the second layer = 256, Kernel size (convolutional window length) = 3, Padding = same, Strides = 1, Neurons in the dense layer = 1, Activation function = Relu, Optimization algorithm = Adam.

Dynamic brittleness index converts to static brittleness index using deep learning

In the following, using three algorithms, show the flowchart the BI_{dyn} converted to BI_{st} (Fig. 7). The workflow shows the transformation of BI_{dyn} to BI_{st} using three algorithms (MLP, LSTM and CNN).

Table 2 shows the parameters used in each of the algorithms

of the deep learning method (MLP, LSTM and CNN) for convert BI_{dyn} to BI_{st}. At first, 24% of the data were separated from the end as blind data, and then, from the other data, 80% were selected as training data and 20% as test data and validation split was selected as 0.1.

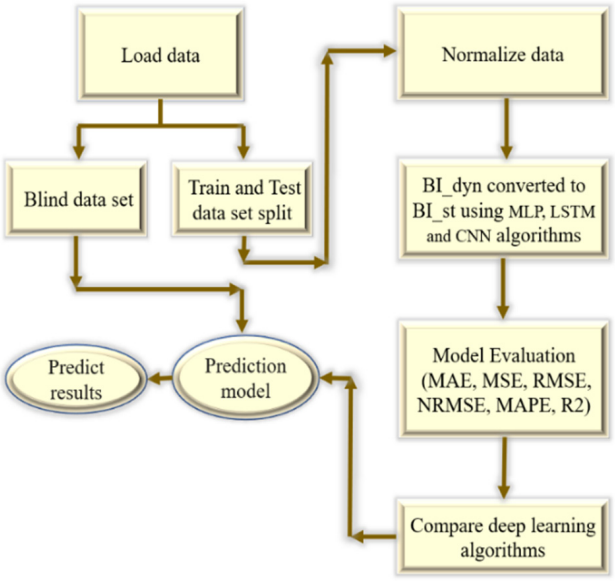

Fig. 7 - Workflow schematic for the transformation of BI_{dyn} to BI_{st} using three algorithms

Static brittleness index using log data and deep learning algorithm

The Empirical relations are usually used to calculate the BI_{st}. In this research, in the previous stage, BI_{st} was estimated from BI_{dyn} using deep learning method. Further, the goal is to estimate the BI_{st} directly from conventional logs. In the introduction, the relation of BI_{st} is introduced (equation 4). BI_{st} was estimated with the conventional data logs of V_p, RHOB, CALIPER, PEF and CGR (the effective features were determined using Auto-encoder's deep learning algorithm) and using MLP, LSTM and CNN algorithms. It has been done, that the training parameters of each of these algorithms are given below. Table 3 shows the parameters used in each of the algorithms of the deep learning algorithm (MLP, LSTM and CNN) for BI_{st} prediction using log data. At first, 24% of the data were separated from the end as blind data, and then, from the other data, 80% were selected as training data and 20% as test data and validation split was selected as 0.1.

MODEL EVALUATION

The performance of the deep learning models for brittleness index prediction are conducted by calculating widely used statistical measures as expressed in equations. 4, 5, 6, 7, 8 and 9. Here, mean absolute percentage error (MAPE), mean absolute error (MAE), Mean Squared Error (MSE), root mean square error (RMSE), Normalized Root Mean Squared Error (NMSRE), and coefficient of determination (R^2) were used to evaluate the performance of model prediction.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Z_{mesured} - Z_{predict})^2 \quad (4)$$

$$RMSE = \sqrt{MSE} \quad (5)$$

$$NRMSE = \frac{RMSE}{MAX(Z_{mesured}) - MIN(Z_{mesured})} \quad (6)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Z_{mesured} - Z_{predict}\right| \quad (7)$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{Z_{mesured} - Z_{predict}}{Z_{mesured}}\right| \quad (8)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Z_{mesured} - Z_{predict})^2}{\sum_{i=1}^{n}(Z_{mesured} - Z_{average})^2} = 1 - \frac{MSE}{\sigma^2} \quad (9)$$

Parameter / Models	Bath size	learning rate	Iteration	Optimization function	Number of layers	Other Description
MLP	100	0.001	300	Adam	2	First hidden layer= 500 nodes Second hidden layer=100 nodes Activation function=Relu
LSTM	50	0.001	300	Adam	2	First hidden layer=200 nodes Second hidden layer=100 nodes Dropout=0.2
CNN	100	0.001	200	Adam	2	First layer the number of filters=128 Second layer the number of filters=256 Kernel size=3, Padding= same Activation function=Relu, Strides=1

Tab. 2 - The same parameters used in each of the algorithms for convert BI_{dyn} to BI_{st}

Parameter / Models	Bath size	learning rate	Iteration	Optimization function	Number of layers	Other Description
MLP	100	0.0001	200	Adam	2	First hidden layer= 500 nodes Second hidden layer=100 nodes Activation function=Relu
LSTM	100	0.001	200	Adam	2	First hidden layer=200 nodes Second hidden layer=100 nodes Dropout=0.2
CNN	100	0.0001	200	Adam	2	First layer the number of filters=128 Second layer the number of filters=256 Kernel size=3, Padding= same Activation function=Relu, Strides=1

Tab. 3 - The same parameters used in each of the algorithms for BI_{st} prediction with log data and three algorithms

RESULTS AND DISCUSSION

Dynamic brittleness index using deep learning method

Tables 4, display BI_{dyn} prediction with log data and deep learning algorithms based on the training (80%) subsets. According to Table 4, for BI_{dyn} training data, three algorithms have a low error, where the MSE values are equal to MSE_{MLP}=1.3720, MSE_{LSTM}=1.7972, MSE_{CNN}=1.5102, and MAE values are equal to MAE_{MLP}=0.5835, MAE_{LSTM}=0.6721, MAE_{CNN}=0.6002, and $RMSE_{MLP}$=1.1713, $RMSE_{LSTM}$=1.3406, $RMSE_{CNN}$=1.2289 respectively.

EP / Models	MLP	LSTM	CNN
MAE	0.5835	0.6721	0.6002
MAPE	1.3209	1.5540	1.3602
MSE	1.3720	1.7972	1.5102
RMSE	1.1713	1.3406	1.2289
NRMSE	0.0332	0.0380	0.0348
R^2	0.9755	0.9679	0.9730

Tab. 4 - BI_{dyn} Prediction errors for training data records using three algorithms

Table 5 displays the BI_{dyn} prediction errors (EP) and coefficient of determination based on the test (20%) subset for three algorithms. Figures 8 provide a comparison for BI_{dyn}measured and BI_{dyn} predicted using three algorithms for train and test data. (Fig. 8a), comparison of BI_{dyn} measured and BI_{dyn} predicted for train and test data using MLP algorithm. (Fig. 8b), BI_{dyn} prediction using LSTM algorithm. (Fig. 8c), BI_{dyn} prediction using CNN algorithm. Blue log (measured BI_{dyn} for training (original data)), orange log (predicted BI_{dyn} for training data), green log (measured BI_{dyn} for test data (original data)), red log (predicted BI_{dyn} for test data). According to Fig. 8 and Table 4 and 5, for BI_{dyn} train and test data, three algorithms MLP, LSTM and CNN have a low error and high coefficient of determination, where the R^2 values for train data are equal to R^2_{MLP}=0.9755, R^2_{LSTM}=0.9679, R^2_{CNN}=0.9730, and for test data MSE values are equal to MSE_{MLP}=0.8867, MSE_{LSTM}=0.4438, MSE_{CNN}=0.2052, and MAE values are equal to MAE_{MLP}=0.7463, MAE_{LSTM}=0.5479, MAE_{CNN}=0.3428, and $RMSE_{MLP}$=0.9416, $RMSE_{LSTM}$=0.6620, $RMSE_{CNN}$=0.4530, and R^2 values are equal to R^2_{MLP}=0.9828, R^2_{LSTM}=0.9914, R^2_{CNN}=0.9960, respectively.

EP / Models	MLP	LSTM	CNN
MAE	0.7463	0.5479	0.3428
MAPE	1.7211	1.3299	0.8131
MSE	0.8867	0.4438	0.2052
RMSE	0.9416	0.6620	0.4530
NRMSE	0.0379	0.0268	0.0182
R^2	0.9828	0.9914	0.9960

Tab. 5 - BI_{dyn} Prediction errors and coefficient of determination for test data

To validate the algorithms, a part of the data was separated from the beginning called blind data. At this stage, the validation of the algorithms has been applied to previously unknown data. Table 6 depicts the predicted BI_{dyn} errors and coefficient of determination based on the blind subsets of data, selected from the 400 data records. To estimate BI_{dyn}, error values for blind data are equal to MSE_{MLP}=5.1687, MSE_{LSTM}=1.7135, MSE_{CNN}=1.0292 and MSE_{MLP}=1.1713, $RMSE_{LSTM}$=1.3406, $RMSE_{CNN}$=1.2289 and $RMSE_{MLP}$=2.2735, $RMSE_{LSTM}$=1.3090, $RMSE_{CNN}$=1.0144, determination coefficient values are equal to R^2_{MLP}=0.9199, R^2_{LSTM}=0.9734, R^2_{CNN}=0.9841. Figure 9 shows a comparison of the predicted BI_{dyn} using three algorithms with the measured BI_{dyn} values for blind data. (Fig. 9a), comparison of BI_{dyn} measured and BI_{dyn} predicted for blind data using MLP algorithm. (Fig. 9b), BI_{dyn} prediction using LSTM algorithm. (Fig. 9c), BI_{dyn} prediction using CNN algorithm (Blue log (BI_{dyn} measured), orange log (BI_{dyn} predicted)). According to the values of table 6 and considering the prediction errors and the coefficient of determination, for blind data in three algorithms, it can be concluded that the most robust and best method of BI_{dyn} prediction is the CNN algorithm and the weakest performance for this case related to the MLP algorithm.

EP \ Models	MLP	LSTM	CNN
MAE	2.0022	1.0905	0.7701
MAPE	4.5768	2.5401	1.6771
MSE	5.1687	1.7135	1.0292
RMSE	2.2735	1.3090	1.0144
NRMSE	0.0645	0.0371	0.0287
R^2	0.9199	0.9734	0.9841

Tab. 6 - BI_{dyn} prediction errors for blind data records using three algorithms

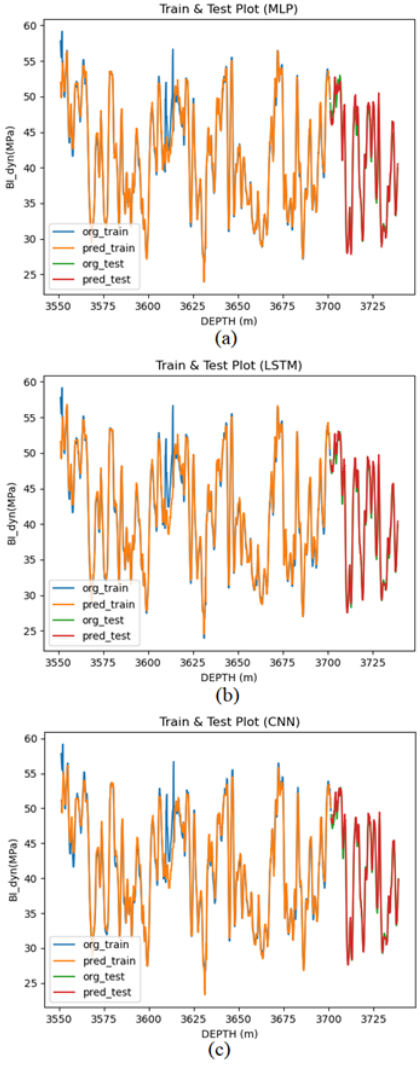

Fig. 8 - Comparison of the predicted BI_{dyn} using three algorithms with the measured BI_{dyn} for train and test data. (a), MLP algorithm. (b), LSTM algorithm. (c), CNN algorithm. Blue log (measured BI_{dyn} for training (original data)), orange log (predicted BI_{dyn} for training data), green log (measured BI_{dyn} for test data (original data)), red log (predicted BI_{dyn} for test data)

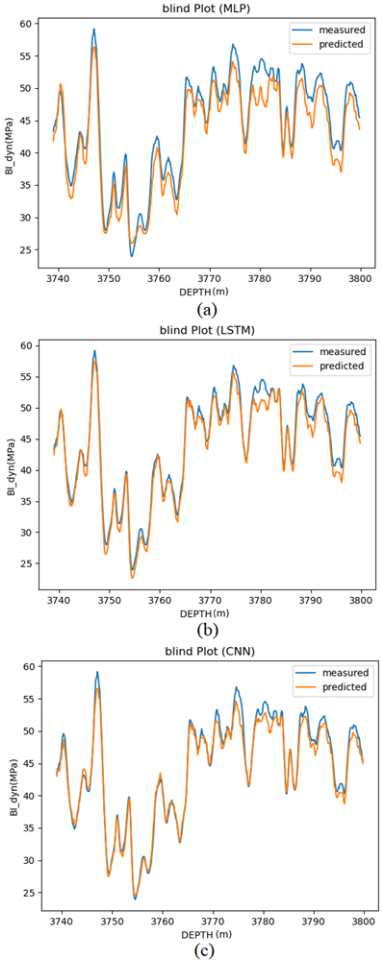

Fig. 9 - Comparison of the predicted BI_{dyn} values using three (MLP, LSTM and CNN) deep learning algorithms with the measured BI_{dyn} values of the blind dataset. (a), BI_{dyn} prediction using MLP. (b), BI_{dyn} prediction using LSTM. (c), BI_{dyn} prediction using CNN

Dynamic brittleness index converts to static brittleness index using deep learning

In the next step, the BI_{dyn} is converted to the BI_{st} using the deep learning algorithms. BI_{st} is calculated from equation 4. The tensile strength is assumed to be 0.1 uniaxial compressive strength. In this research, 12 laboratory samples of UCS have been available (Table 1). In order to generalize these samples to the entire target range, we first calculate the UCS with the Christaras (Christaras *et alii*, 1997) relations (equation 12) at the depths where the laboratory samples of UCS were recorded, and then the regression between the UCS obtained from the log and we determine the UCS calculated in the laboratory, and the obtained regression is equal to 0.98 with a 2nd degree equation and we generalize it to the entire depth range. Figure 10 shows the regression and equation obtained between the uniaxial

compressive strength of the log and the uniaxial compressive strength of the laboratory samples.

$$UCS = 9.95V_p^{1.21} \quad (10)$$

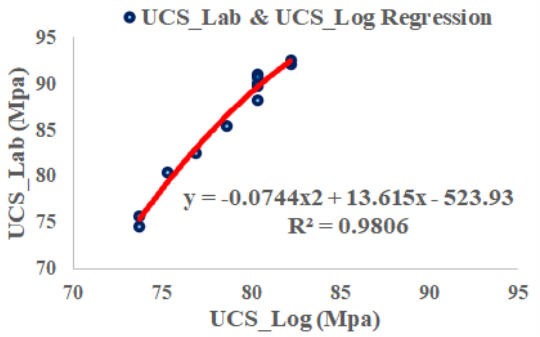

Fig. 10 - Workflow schematic for the transformation of BI_{dyn} to BI_{st} using three algorithms

Tables 7, display the model evaluation parameters, the transformation of the BI_{dyn} to the BI_{st} for the training data. According to Table 7, for the transformation of the BIdyn to the BI_{st} training data, three algorithms have a low error, where the MSE and values are equal to MSE_{MLP}=1.0031, MSE_{LSTM}=0.8393, MSE_{CNN}=0.8855, and MAE values are equal to MAE_{MLP}=0.6482, MAE_{LSTM}=0.5409, MAE_{CNN}=0.5716, and $RMSE_{MLP}$=1.0015, $RMSE_{LSTM}$=0.9161, $RMSE_{CNN}$=0.9410 respectively.

EP \ Models	MLP	LSTM	CNN
MAE	0.6482	0.5409	0.5716
MAPE	8.5655	6.6329	6.6202
MSE	1.0031	0.8393	0.8855
RMSE	1.0015	0.9161	0.9410
NRMSE	0.0466	0.0426	0.0438
R^2	0.9575	0.9644	0.9652

Tab. 7 - Model evaluation parameters, BI_{dyn} convert to BI_{st} for training data

Table 8 display the model evaluation parameters, the transformation of the BI_{dyn} to the BI_{st} for the test (20%) subset with three algorithms. According to Table 8, for the transformation of the BI_{dyn} to the BI_{st} training data, three algorithms have a low error, where the MSE and values are equal to MSE_{MLP}=0.8909, MSE_{LSTM}=0.8351, MSE_{CNN}=0.4273, and MAE values are equal to MAE_{MLP}=0.6891, MAE_{LSTM}=0.6658, MAE_{CNN}=0.5560, and RMSE values are equal to $RMSE_{MLP}$=0.9439, $RMSE_{LSTM}$=0.9138, $RMSE_{CNN}$=0.6537 respectively. Figures 11 provide a comparison for BI_{st} measured and BI_{st} predicted using three algorithms for train and test data. (Fig. 11a), comparison of BI_{st} measured

EP \ Models	MLP	LSTM	CNN
MAE	0.6891	0.6658	0.5560
MAPE	6.3602	5.1927	4.3418
MSE	0.8909	0.8351	0.4273
RMSE	0.9439	0.9138	0.6537
NRMSE	0.0509	0.0492	0.0432
R^2	0.9713	0.9731	0.9797

Tab. 8 - Model evaluation parameters, BI_{dyn} convert to BI_{st} for test data

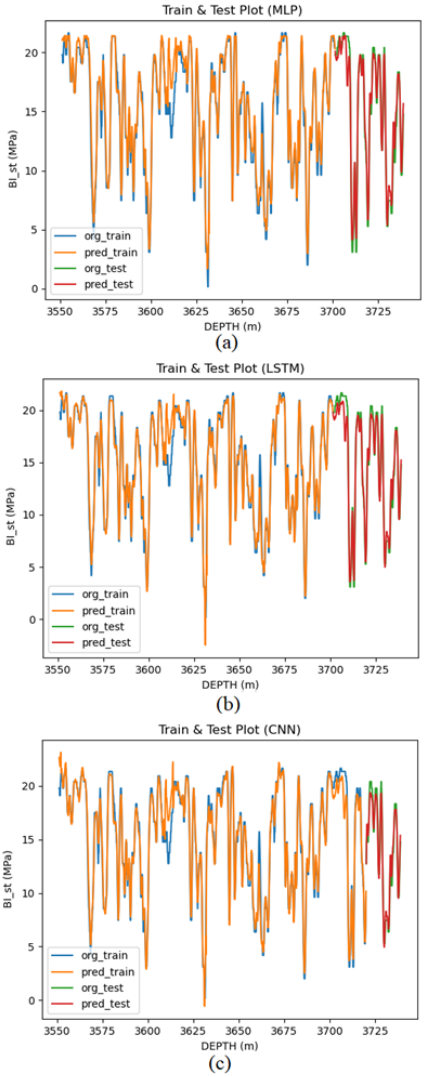

Fig. 11 - Display of BI_{st} prediction from BI_{dyn} with three algorithms for train and test data. (a), BI_{st} prediction with MLP algorithm. (b), BI_{st} prediction with LSTM. (c), BI_{st} prediction with CNN. Blue log (BI_{st} measured for training (original data)), orange log (BI_{st} predicted for training data), green log (BI_{st} measured for test data (original data)), red log (BI_{st} predicted for test data)

and BI_{st} predicted for train and test data using MLP algorithm. (Fig. 11b), BI_{st} prediction using LSTM algorithm. (Fig. 11c), BI_{st} prediction using CNN algorithm. Blue log (measured BI_{st} for training (original data)), orange log (predicted BI_{st} for training data), green log (measured BI_{st} for test data (original data)), red log (predicted BI_{dyn} for test data). According to Fig. 11 and Table 7 and 8, for BI_{st} train and test data, three algorithms MLP, LSTM and CNN have a low error and high coefficient of determination, where the R^2 values for train data are equal to R^2_{MLP}=0.9755, R^2_{LSTM}=0.9644, R^2_{CNN}=0.9652, and R^2 values for test data are equal to R^2_{MLP}=0.9713, R^2_{LSTM}=0.9731, R^2_{CNN}=0.9797, respectively.

Table 9, depicts the transformation of the BI_{dyn} to the BI_{st} errors and coefficient of determination based on the blind subsets of data, selected from the 400 data records. To estimate BI_{st}, error values for blind data are equal to MSE_{MLP}=1.4823, MSE_{LSTM}=0.9425, MSE_{CNN}=0.8444, and MAE values are equal to MAE_{MLP}=1.0832, MAE_{LSTM}=0.7598, MAE_{CNN}=0.7746, and RMSE values are equal to $RMSE_{MLP}$=1.2175, $RMSE_{LSTM}$=0.9708, $RMSE_{CNN}$=0.9189, and determination coefficient values are equal to R^2_{MLP}=0.9457, R^2_{LSTM}=0.9655, R^2_{CNN}=0.9691. Figure 12 shows a comparison of the BI_{st} prediction from BI_{dyn} and BI_{st} measured for blind data for blind data. (Fig. 12a), comparison of BI_{st} measured and BI_{st} predicted for blind data using MLP algorithm. (Fig. 12b), BI_{dyn} prediction using LSTM algorithm. (Fig. 12c), BI_{st} prediction using CNN algorithm (Blue log (BI_{st} measured), orange log (BI_{st} predicted)). According to the values of Table 9 and considering the prediction errors and the coefficient of determination, for blind data in three algorithms, it can be concluded that the most robust and best method of BI_{dyn} is converted to the BI_{st} prediction is the CNN algorithm and the weakest performance for this case related to the MLP algorithm.

EP \ Models	MLP	LSTM	CNN
MAE	1.0832	0.7598	0.7746
MAPE	14.773	8.8215	10.3567
MSE	1.4823	0.9425	0.8444
RMSE	1.2175	0.9708	0.9189
NRMSE	0.0567	0.0452	0.0428
R^2	0.9457	0.9655	0.9691

Tab. 9 - Model evaluation parameters, BI_{dyn} convert to BI_{st} for blind data

Static brittleness index using log data and deep learning algorithm

Table 10, displays BI_{st} prediction errors respectively based on the training (80%) subsets. According to Table 10, for BI_{st} training data, three algorithms have a low error, where the

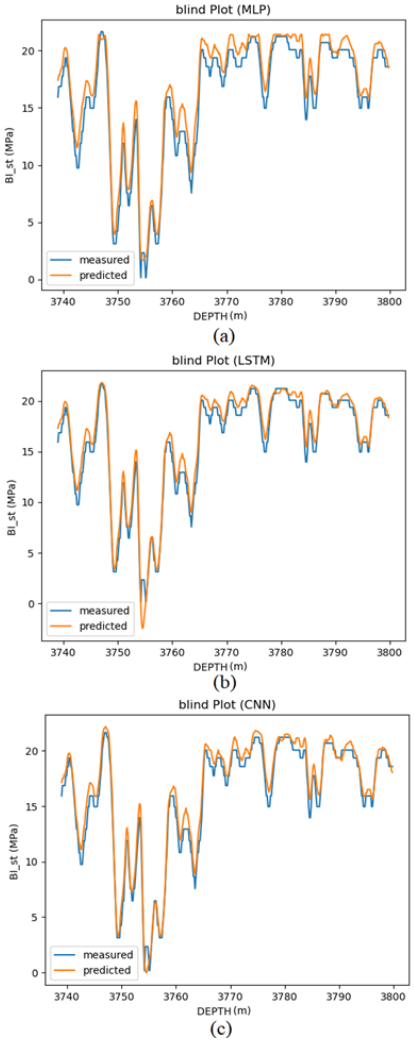

Fig. 12 - Display of BI_{st} prediction from BI_{dyn} with three algorithms for blind data. (a), BI_{st} prediction for blind data with MLP algorithm. (b), BI_{st} prediction with LSTM. (c), BI_{st} prediction with CNN. Blue log (BI_{st} measured), orange log (BI_{st} predicted)

MSE and values are equal to MSE_{MLP}=0.0340, MSE_{LSTM}=0.0102, MSE_{CNN}=0.0007 and MAE values are equal to MAE_{MLP}=0.1177, MAE_{LSTM}=0.0685, MAE_{CNN}=0.0187, and $RMSE_{MLP}$=0.1845,

EP \ Models	MLP	LSTM	CNN
MAE	0.1177	0.0685	0.0187
MAPE	0.8482	0.7058	0.1866
MSE	0.0340	0.0102	0.0007
RMSE	0.1845	0.1010	0.0276
NRMSE	0.0086	0.0047	0.0012
R^2	0.9985	0.9995	0.9999

Tab. 10 - BI_{st} Prediction errors and accuracy for training data records with three algorithms

$RMSE_{LSTM}$=0.1010, $RMSE_{CNN}$=0.0276 respectively.

Table 11 displays the BI_{st} prediction errors (EP) and coefficient of determination based on the test (20%) subset for three algorithms. Figures 13 provides a comparison for BI_{st} measured and BI_{st} predicted using three algorithms for train and test data. (Fig. 13a), comparison of BI_{st} measured and BI_{st} predicted for train and test data using MLP algorithm. (Fig. 13b), BI_{st} prediction using LSTM algorithm. (Fig. 13c), BI_{st} prediction using CNN algorithm. Blue log (measured BI_{st} for training (original data)), orange log (predicted BI_{st} for training data), green log (measured BI_{st} for test data (original data)), red log (predicted BI_{st} for test data). According to Fig. 13 and Table 10 and 11, for BI_{st} train and test data, three algorithms MLP, LSTM and CNN have a low error and high coefficient of determination, where the R^2 values for train data are equal to R^2_{MLP}=0.9985, R^2_{LSTM}=0.9995, R^2_{CNN}=0.9999, and for test data the MSE and values are equal to MSE_{MLP}=0.1290, MSE_{LSTM}=0.0241, MSE_{CNN}=0.0052 and MAE values are equal to MAE_{MLP}=0.2420, MAE_{LSTM}=0.0876, MAE_{CNN}=0.0484, and $RMSE_{MLP}$=0.3592, $RMSE_{LSTM}$=0.1553, $RMSE_{CNN}$=0.0727 and R^2 values are equal to R^2_{MLP}=0.9958, R^2_{LSTM}=0.9902, R^2_{CNN}=0.9998, respectively.

EP \ DL models	MLP	LSTM	CNN
MAE	0.2420	0.0876	0.0484
MAPE	1.5688	0.6840	0.3865
MSE	0.1290	0.0241	0.0052
RMSE	0.3592	0.1553	0.0727
NRMSE	0.0193	0.0083	0.0039
R^2	0.9958	0.9992	0.9998

Tab. 11 - BI_{st} prediction errors for test data records using three algorithms

Table 12 depicts the predicted BI_{st} errors and coefficient of determination based on the blind subsets of data, selected from the 400 data records. To estimate BI_{st}, error values for blind data are equal to MSE_{MLP}=1.2071, MSE_{LSTM}=0.9076, MSE_{CNN}=0.6719 and MAE_{MLP}=0.8980, MAE_{LSTM}=0.7866, MAE_{CNN}=0.6434, and $RMSE_{MLP}$=1.0987, $RMSE_{LSTM}$=0.9527, $RMSE_{CNN}$=0.8197 and determination coefficient values are equal to R^2_{MLP}=0.9557, R^2_{LSTM}=0.9667, R^2_{CNN}=0.9754. Figure 14 shows a comparison of the predicted BI_{st} using three algorithms with the measured BI_{st} values for blind data (Fig. 14a), comparison of BI_{st} measured and BI_{st} predicted for blind data using MLP algorithm. (Fig. 14b), BI_{st} prediction using LSTM algorithm. (Fig. 14c), BI_{st} prediction using CNN algorithm (Blue log (BI_{st} measured), orange log (BI_{st} predicted)). Figure 15 shows the coefficient of determination of the blind data of the BI_{st} measured and the BI_{st} predicted for three algorithms. (Fig. 15a), the coefficient of determination of the blind data using the MLP algorithm. (Fig. 15b), the coefficient of determination using the LSTM algorithm. (Fig. 15c), the

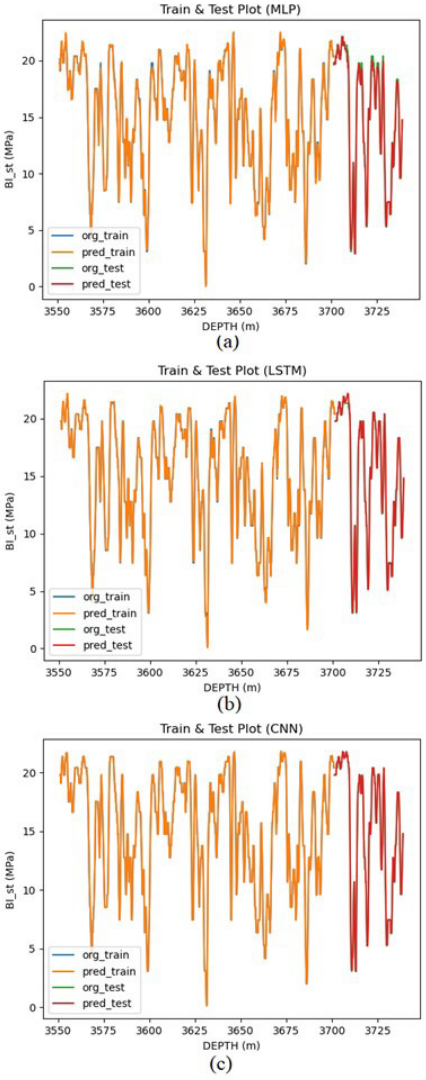

Fig. 13 - Display of BI_{st} predict from log data with three algorithms for train and test data. (a), BI_{st} prediction with MLP algorithm. (b), BI_{st} prediction with LSTM. (c), BI_{st} prediction with CNN. Blue log (BI_{st} measured for training (original data)), orange log (BI_{st} predicted for training data), green log (BI_{st} measured for test data (original data)), red log (BI_{st} predicted for test data)

coefficient of determination using the CNN algorithm. According to the values of Table 12 and considering the prediction errors and the coefficient of determination, for blind data in three algorithms, it can be concluded that the most robust and best method of BI_{st} prediction is the CNN algorithm and the weakest performance for this case related to the MLP algorithm.

Figure 16a shows the comparison of all measured and predicted dynamic brittleness index data BI_{dyn} measured (blue color) and BI_{dyn} predicted (orange color). (Fig. 16b) shows the comparison of all predicted static brittleness index data

EP \ Models	MLP	LSTM	CNN
MAE	0.8980	0.7866	0.6434
MAPE	17.5282	12.6217	14.1078
MSE	1.2071	0.9076	0.6719
RMSE	1.0987	0.9527	0.8197
NRMSE	0.0511	0.0443	0.0381
R^2	0.9557	0.9667	0.9754

Tab. 12 - BI_{st} prediction errors for blind data with three algorithms

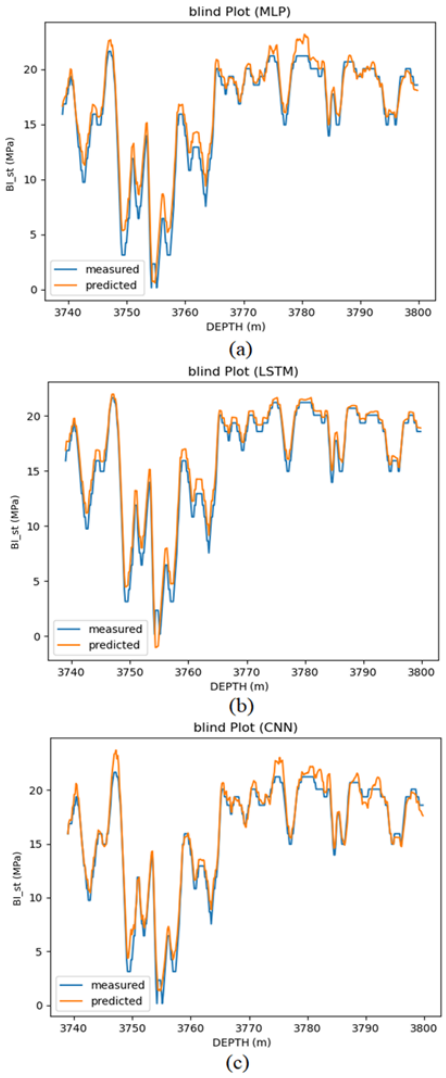

Fig. 14 - Display of BI_{st} predict from log data with three algorithms for blind data. (a), BI_{st} prediction with MLP. (b), BI_{st} prediction with LSTM. (c), BI_{st} prediction with CNN. Blue log (BI_{st} measured), orange log (BI_{st} predicted)

(blue color) and core data (orange color). (Fig. 16c) shows the enlargement of the comparison between core samples and

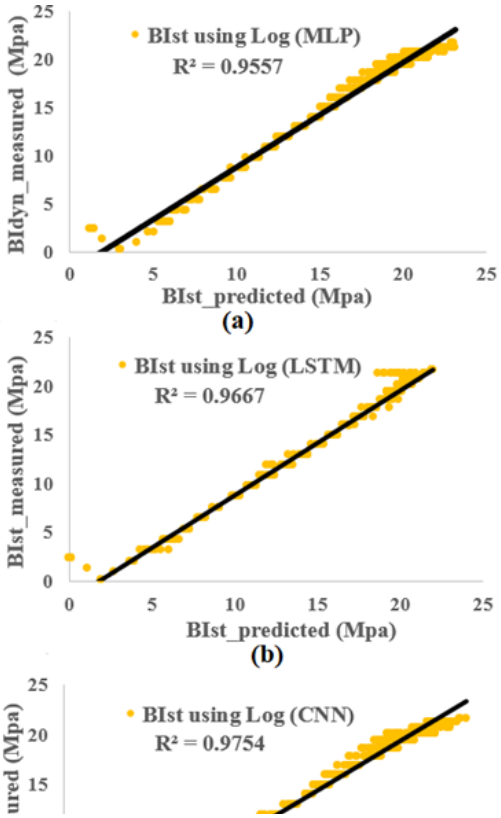

Fig. 15 - Display of coefficient of determination of blind data for BI_{st} measured and BI_{st} predicted with three deep algorithms. (a), R^2 using MLP algorithm. (b), R^2 using LSTM algorithm. (c), R^2 using CNN algorithm

predicted brittleness index data. (Fig. 16d) shows the brittleness index obtained with the volume percentage of minerals with equation 5. In (Fig. 16d), sandstone ranges have a higher brittleness index than carbonate ranges. According to Fig. 16, there is a very good match between core data and predicted static brittleness index. Also, according to the comparison between the index obtained from the volume percentage of minerals and the obtained static and dynamic brittleness index, there is a good match. Also, according to relation 5 and also the brittleness index calculated from lithology and lithology column, it can be stated that in the carbonate environment, in the areas where the lithology is limestone and dolomite, it has a lower brittleness index compared to the areas where the lithology is sandstone.

The deep learning algorithm is one of the new and high-accuracy methods for predicting BI. In this paper, three deep learning algorithms (MLP, LSTM and CNN) are used to predict BI_{dyn} and BI_{st}. In this research, the aim is to check the accuracy of the algorithms used in predicting the BI_{dyn} and BI_{st} of the blind data, that these algorithms can be used in wells where we do not have an actual data. The results show the high accuracy of LSTM and CNN algorithm for BI_{dyn} and BI_{st} prediction.

Therefore, the accuracy and robustness of the prediction results of these techniques have more advantages than MLP model and traditional empirical models.

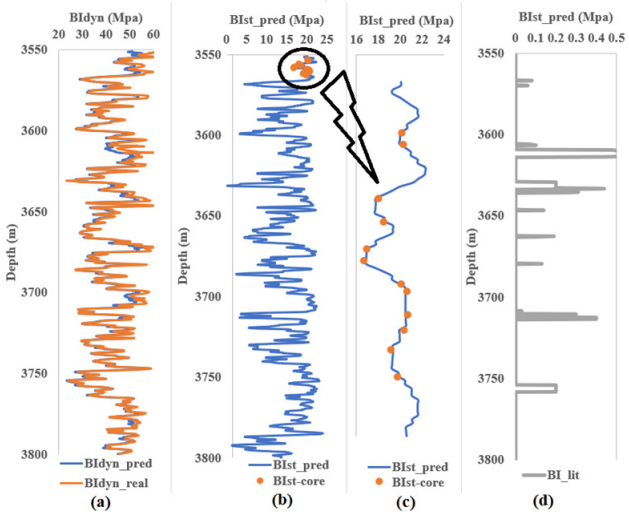

Fig. 16 - BI_{dyn} measured (blue color) and BI_{dyn} predicted (orange color). (b), BI_{st} predicted (blue color) and core data (orange color). (c), the enlargement of the comparison between core samples and BI_{st} predicted. (d), brittleness index with the volume percentage of minerals

CONCLUSION

One of the important stages of geomechanical modeling is the determination of the BI, which is mainly determined by empirical relations. Actual data of well from one of Iran's hydrocarbon fields is selected in this study to perform relevant tests. In this study, some models are stablished for predicting BI_{dyn} and BI_{st} values based on three MLP, LSTM and CNN algorithms. To achieve the goals, first, the Auto-encoder algorithm was used to select the effective features, and the effective features of the Vp, RHOB, CALIPER, PEF and CGR were selected for BI predicted. To estimate the BI of the structure of the MLP model including two layers where the first layer is 500 nodes and the second layer is 100 nodes, the LSTM model including two layers where the first layer is 200 nodes and the second layer is 100 nodes and the dropout is 0.2, and for CNN model consists of two layers, 128 filters of the first layer, 256 filters of the second layer were selected. In this research, the BI_{dyn} was first estimated with three algorithms and conventional data logs, and then two methods were used to estimate the BI_{st}. The BI_{dyn} was estimated with log data and three algorithms, and then three algorithms were used to convert BI_{dyn} to BI_{st}, in the following, instead of using two steps to estimate the BI_{st}, this parameter was estimated directly from conventional logs using three algorithms (MLP, LSTM and CNN). To ensure the results of the algorithms, the evaluation of the models was done with the parameters of MAE, MAPE, MSE, RMSE and R^2.

In the next step, the three models were applied on the blind data of the BI_{dyn} and BI_{st}, R^2 values obtained for BI_{dyn}; R^2_{MLP}=0.9199, R^2_{LSTM}=0.9734, R^2_{CNN}=0.9841 and R^2 values obtained for BI_{dyn} converted to BI_{st}; R^2_{MLP}=0.9457, R^2_{LSTM}=0.9655, R^2_{CNN}=0.9691 and R^2 values obtained for BI_{st}; R^2_{MLP}=0.9557, R^2_{LSTM}=0.9667, R^2_{CNN}=0.9754. In continue according to the comparison between the index obtained from the volume percentage of minerals and the obtained static and dynamic brittleness index, there is a good match. Moreover, the BI_{dyn} and BI_{st} are predicted by the proposed method and traditional empirical models, it has been demonstrated that the introduced method, as an efficient deep learning model, outperforms MLP models and traditional empirical models in their prediction accuracy and robustness. While MLP network achieved relatively satisfactory results in BI_{dyn} and BI_{st} prediction, but compared to the LSTM and CNN algorithms, it shows less accuracy and more error. the results of which show the robust of deep learning algorithms in BI_{dyn} and BI_{st} prediction.

ABBREVIATIONS

AI	Artificial intelligence
ANN	Artificial neural network
BI_{dyn}	Dynamic Brittleness Index
BI_{st}	Static Brittleness Index
CGR	Corrected Gamma Ray
CHAL	Caliper log
CNL	Compensate neutron log
CNN	Convolutional neural network
DL	Deep learning
DNN	Deep neural network
DT	Acoustic (sonic) log
ELM	Extreme learning machine
ENN	Elman neural network
EP	Evaluation Parameter
PEF	Photoelectric Log
FFANN	Feedforward artificial neural network
FL	Fuzzy logic
GA	Genetic algorithm
GEP	Gene expression programming
GR	Gamma ray log
GRNN	General regression neural network
LSSVM	Least-squares support-vector machines
LSTM	Long short-term memory networks
MAE	Mean absolute error
MAPE	Mean absolute percentage error
MF	Memetic firefly
ML	Machine learning technique
MLEM	Multi extreme learning machine
MLP	Multi-layer perceptron
MSE	Mean square error
NRMSE	Normalized Root Mean Squared Error

R^2	Coefficient of determination	SGR	Sum Gamma Ray
RHOB	Density log	SVM	Support vector machine
RMSE	Root mean square error	SVR	Support vector regression
RNN	Recurrent neutral network	UCS	Uniaxial Compressive Strength
RS	Shallow lateral resistivity log	V_s	Shear wave velocity
RT	Formation true resistivity	V_p	Compressional wave velocity

REFERENCES

Abadi M., Barham P., Chen J., Chen Z., Davis A., Dean J., Devin M., Ghemawat S., Irving G., Isard M. & Kudlur M. (2016) - *Tensorflow: a system for large-scale machine learning.* In: 12th {USENIX} Symposium on Operating Systems Design and Implementation (OSDI): 265-283.

Alavi A.H., Gandomi A.H. Mollahasani A. Heshmati A.A.R. & Rashed A. (2010) - *Modeling of maximum dry density and optimum moisture content of stabilized soil using artificial neural networks.* Journal of Plant Nutrition and Soil Science, **173**(3): 368-379. https://doi.org/10.1002/jpln.200800233.

Alsina E.F., Chica M., Trawinski K. & Regattieri A. (2017) - *On the use of machine learning methods to predict component reliability from data-driven industrial case studies.* Int. J. Adv. Manuf. Technol, **94**: 2419-2433. https://doi.org/10.1007/s00170-017-1039-x.

Altindag R. (2010) - *Assessment of some brittleness indexes in rock-drilling efficiency.* Rock mechanics and rock engineering, **43**(3): 361-370. https://doi:10.1007/s00603-009-0057-x.

Altindag R. (2003) - *Correlation of specific energy with rock brittleness concepts on rock cutting.* J. South Afr. Inst. Min. Metall., **103**(4): 163-171.

Anemangely M., Ramezanzadeh A., Amiri H. & Hoseinpour S.A. (2018) - *Machine learning technique for the prediction of shear wave velocity using petrophysical logs*. J. Pet. Sci. Eng., **174**. https://doi:10.1016/j.petrol.2018.11.032.

Asemi F., Lakirouhani A., Nicksiar M. & Zohdi A. (2024) - *A New Rock Brittleness Index Based on Crack Initiation and Crack Damage Stress Thresholds.* International Journal of Geomechanics, **24**. https://doi.org/10.1061/IJGNAI.GMENG-9151.

Bengio Y., Simard P. & Frasconi P. (1994) - *Learning long-term dependencies with gradient descent is difficult.* IEEE Trans. Neural Network, **5**(2): 157-166.

Chollet F. (2015) - *Keras*. https://github.com/fchollet/keras.

Christaras B. (1997) - *Landslides in iliolitic and marly formations. Examples from north-westem Greece.* Engineering Geology, **47**(1-2): 57-69.

Consonni V., Ballabio D. & Todeschini R. (2010) - *Evaluation of model predictive ability by external validation techniques.* J. Chemometr., **24**: 194-201. https://doi.org/10.1002/cem.1290.

Duchi J., Hazan E. & Singer Y. (2011) - *Adaptive subgradient methods for online learning and stochastic optimization.* J.Machine Learning Research., **12**: 2121-2159.

Da Silva M. (2013) - *Production correlation to 3D seismic attributes in the Barnett Shale, Texas*: M.S. thesis, The University of Oklahoma.

Fang Z., Yunjie Z., Yu L., Jixiang H., Xuechun Z. & Yaoli Sh. (2024) - *A seismic prediction method of reservoir brittleness based on mineral composition and pore structure*. Journals earth-science, **11**. https://doi.org/10.3389/feart.2023.1326861.

Ghobadi M.H. & Naseri F. (2016) - *Rock Brittleness Prediction Using Geomechanical Properties of Hamekasi Limestone: Regression and Artificial Neural Networks Analysis.* JGeope **6**(1): 19-33.

Ghobadi M.H., Amiri M. & Rahimi Shahid M. (2023) - *The estimation of Brittleness indexes of Qom Formation sandstones in northern Hamedan using the ratio between point load index and porosity.* New finding in applied geology. https://doi:10.22084/NFAG.2022.25272.1493.

Goktan R.M. & Yilmaz N.G. (2005) - *A new methodology for the analysis of the relationship between rock brittleness index and drag pick cutting.* J. South Afr. Inst. Min. Metall., **105**: 727-733.

Guktan R.M. (1991) - *Brittleness and micro-scale rock cutting efficiency.* Mining Science and Technology, **13**(3): 237-241.

Greff K., Srivastava R.K., Koutník J., Steunebrink B.R. & Schmidhuber J. (2017) - *LSTM: A Search Space Odyssey.* Ieee Transactions on Neural Networks and Learning Systems, **28**(10): 2222-2232. https://doi.org/10.48550/arXiv.1503.04069.

Gu J. (2018) - *Recent advances in convolutional neural networks.* Pattern Recognition, **77**: 354-377. https://doi:10.1016/j.patcog.10.013.

Guo Y., Liu Y. Oerlemans A., Lao S., Wu S. & Lew M.S. (2015) - *Deep learning for visual understanding: A review. Neurocomputing,* **187**: 27-48. https://doi: 10.1016/j.neucom.2015.09.116.

Hetenyi M. (1966) - *Handbook of experimental stress analysis*, Wiley, New York, 15.

Hochreiter S. & Schmidhuber J. (1997) - *Long short-term memory.* Neural Comput.. **9**(8): 1735-1780.

Hucka V. & Das B. (1974) - *Brittleness determination of rocks by different methods.* In International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts, **11**(10): 389-392. Pergamon.

Hunter J.D. (2007) - *Matplotlib: a 2D graphics environment.* Comput. Sci. Eng. **9**(3): 90-95.

Jarvie D.M., Hill R.J. Ruble T.E. & Pollastro R.M. (2007) - *Unconventional shale-gas systems: the Mississippian Barnett Shale of North-Central Texas as one model for thermogenic shale-gas assessment.* AAPG Bull., **91**(4): 475-99. https://doi.org/10.1306/12190606068.

Jin X., Shah S., Roegiers J. & Zhang B. (2014) - *Fracability evaluation in shale reservoirs - An integrated petrophysics and geomechanics approach*. Presented at the SPE Annual Technical Conference and Exhibition.

Kahramana S. (2002) - *Correlation of TBM and drilling machine performances with rock brittleness.* Eng. Geol. **65**(4): 269-283.

Kelishami S.B.A., Rezaei M. & Mohebian R. (2022) - *A new approach to estimate and delineate the geothermal gradient of Iran.* Geothermics, **103**: 102428.

Kingm D.P. & Ba J. (2014) - *Adam: A Method for Stochastic Optimization.* https://arxiv. org/abs/1412.6980.

Lawn B. & Marshall D.B. (1979) - *Hardness, toughness, and brittleness: an indentation Analysis.* Journal of the American ceramic society, **62**(7-8): 347-350.

Li Z.H., Zhang S.Y. & Cao F. (2020) - *Brittleness Evaluation of Chang7 Reservoir in Longdong Area Based on Mechanical Parameters and Stress.* IOP Conference Series: Earth and Environmental Science 514.

Lecun Y., Bottou L., Bengio L. & Haffner P. (1998) - *Gradient -based learning applied to document recognition*. Proceedings of the IEEE, **86**(11): 2278-2324. https://doi: 10.1109/5.72679.

Mohebian R., Riahi M.A. & Kadkhodaie-Ilkhchi A. (2017) - *A comparative study of the neural network, fuzzy logic, and nero-fuzzy systems in seismic reservoir characterization: an example from arab (surmeh) reservoir as an Iranian gas field, Persian Gulf basin*. Iranian Journal of Oil and Gas Science and Technology, **6**(4): 33-55. https://doi.org/10.22050/ijogst.2017.53907.

Mohebian R. & Riahi M.A. (2019) - *Integrating neural, fuzzy logic, and neuro-fuzzy approaches using Ant Colony Optimisation for continuous domains to determine carbonate reservoir facies.* Bollettino di Geofisica Teorica ed Applicata, **60**(4). https://doi: 10.4430/bgta0281.

Mohebian R., Bagheri H., Kheirollahi M. & Bahrami H. (2021) - *Permeability estimation using an integration of multi-resolution graph-based clustering and rock typing methods in an Iranian Carbonate Reservoir.* Journal of Petroleum Science and Technology, **11**(3): 49. https://doi:10.22078/JPST.2022.4737.1785.

Morley A. (1954) - *Strength of materials.* 11th ed. Longmans, Green, London, 532.

Obert L. & Duvall W.I. (1967) - *Rock mechanics and the design of structures in rock.* (No. BOOK). J. Wiley.

Oliphant T.E. (2006) - *A guide to NumPy.* Methods **1**: 85.

Ore T. & Gao D. (2021) - *Supervised machine learning to predict brittleness using well logs and seismic signal attributes: Methods and application in an unconventional reservoir.* Applied Geoscience & Energy. https://doi:10.1190/segam2021-3594773.1.

Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau D., Brucher M., Perrot M. & Duchesnay E. (2011) - *Scikit-learn: machine learning in Python.* J. Mach. Learn. Res. ,**12**: 2825-2830.

Perez Altamar R. (2013) - *Brittleness estimation from seismic measurements in unconventional reservoirs: Application to the Barnett Shale*. Ph.D. dissertation, The University of Oklahoma.

Qiuan R.K., Liu T., Liu J.Z., Liu X.W., He Z.L. & Jiang D.J. (2019) - *Construction of a novel brittleness index equation and analysis of anisotropic brittleness characteristics for unconventional shale formations*. Petroleum Science. https://doi.org/10.1007/s12182-019-00372-6.

Rahimi Shahid M., Lashkaripour Gh. R. & Hafezi Moghaddas N. (2023) - *Determining the brittleness index of limestones in western Iran (Zone 38) using experimental relationships*. The 42st National Geosciences Congress.

Ramsay J.G. (1967) - *Folding and fracturing of rocks.* McGraw Hill Book Company, 568.

Rickman R.M., Mullen J., Petre W., Grieser D. & Kundert A. (2008) - *A practical use of shale petrophysics for stimulation design optimization: All shale plays are not clones of the Barnett Shale*. Presented at the SPE Annual Technical Conference and Exhibition.

Shi X., Liu G., Cheng Y., Yang L., Jiang H., Chen L., Jiang S.H. & Wang J. (2016) - *Brittleness index prediction in shale gas reservoirs based on efficient network models.* Journal of Natural Gas Science and Engineering. http://dx.doi.org/10.1016/j.jngse.2016.09.009.

Tieleman T. & Hinton G. (2012) - *Lecture 6.5-rmsprop: divide the gradient by a running average of its recent magnitude.* COURSERA: Neural Networks for Machine Learning **4**: 26-30.

Valim S.M. & Antia L.S. (2021) - *The Use of Well-Log Data in the Geomechanical Characterization of Middle Cambrian Tight Sandstone Formation: A Case Study from Eastern Pomerania, Poland.* https://doi.org/10.3390/en14196022.

Wang F.P. & Gale J.F. (2009) - *Screening criteria for shale-gas systems*. Gulf Coast Assoc. Geol. Soc. Trans., **59**: 779-793.

Wang D.B., Ge H.K., Wang X.Q., Wang J.B., Meng F.B., Suo Y. & Han P. (2015) - *A novel experimental approach for fracability evaluation in tight-gas reservoirs.* J. Nat. Gas Sci. Eng. **23**: 239-249.

Wang T.T., Zhang H.T. & Ranjith P.G. (2021) - *Evaluation and analysis of sandstone brittleness under the influence of temperature.* Geomechanics and Geophysics for Geo - Energy and Geo - Resources, **8**.

Wang H., He M., Zhao J. & Zhang Y. (2023) - *Digital drilling-based determination of rock anisotropy and anisotropic effect on cutter wear*. Quarterly Journal of Engineering Geology and Hydrogeology, **56**. https://doi.org/10.1144/qjegh2022-103.

Wang H., He M., Zhao J., Zhang Y. & Yang B. (2023) - *Cutting energy characteristics for brittleness evaluation of rock using digital drilling method. Engineering Geology*. **319**. https://doi.org/10.1016/j.enggeo.2023.107099.

Wang J., Du J., Li W., Chen X., Zhang H., Tao W., Li Z. & Hao R. (2023) - *Brittleness index evaluation of gas-bearing sandstone under triaxial compression conditions.* Geomechanics and Geophysics for Geo-Energy and Geo-Resources. https://doi.org/10.1007/s40948-023-00713-1.

Xiao W., Zhang D. & Yang H. (2020) - *The evaluation criteria for rock brittleness based on double-body analysis under uniaxial compression.* Geomechanics and Geophysics for Geo - Energy and Geo - Resources **6**: 61-66.

Xie H.P., Ju Y. & Li L.Y. (2008) - Energy mechanism of deformation and failure of rock masses. Chinese Journal of Rock Mechanics and Engineering 27: 1729-1740.

Xu J.C., Pu H. & Sha Z.H. (2021) - *Experimental Study on the Effect of Brittleness on the Dynamic Mechanical Behaviors of the Coal Measures Sandstone.* Advances in Civil Engineering.

Yagiz S. (2009) - *Assessment of brittleness using rock strength and density with punch penetration test.* Tunnelling and Underground Space Technology, **24**(1): 66-74.

Yoo H.J. (2015) - *Deep Convolution Neural Networks in Computer Vision: a Review*. IEIE Transactions on Smart Processing and Computing, **4**: 35-43. https://doi: 10.5573/IEIESPC.2015.4.1.035.

Zhang B.T., Zhao J. & Marfurt K.J. (2014b) - *Horizonbased semi-automated nonhyperbolic velocity analysis*. Geophysics, **79**(6): U15-U23. https://doi: 10.1190/geo2014- 0112.1.

Zhang D.C., Ranjith P.G. & Perera MSA. (2016) - *The brittleness indices used in rock mechanics and their application in shale hydraulic fracturing: a review.* J Pet. Sci. Eng., **143**: 158-170.

Zhang L.G., Qu S.N. & Li S.B. (2017) - *The properties and relationship of the brittleness, mineral composition and porosity for tight oil reservoirs. Strength*. Fracture and Complexity **10**: 89-99.

Zhang X., Xu J., Shaikh F., Sun L. & Cao Y. (2022) - *Rock Brittleness Evaluation Index based on Ultimate Elastic Strain Energy.* Research square. https://doi.org/10.21203/rs.3.rs-1688680/v1.

Zhang T., Lin K., Wen X., Zhang Y. & Zhao L. (2024) - *Direct inversion of brittleness index in time-frequency mixed domain.* Geoenergy science and engineering, **238**. https://doi.org/10.1016/j.geoen.2024.212920.

Received February 2024 - Accepted June 2024