Proceedings of 8th International Conference on Mechanical Engineering and Renewable Energy 2025
10-12 December, 2025, Chattogram, BANGLADESH

SciEn Conference Series: Engineering Vol. 4, 2026, pp 549-554 https://doi.org/10.38032/scse.2026.4.239

Explainable AI-Based Geophysical Logging Parameters Selection for Enhancing Rock Brittleness Prediction

Mohammad Islam Miah[1,*], *Md. Maruf Alam*[2], *Travis Wiens*[1]

[1]Department of Mechanical Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada.
[2]Instituite of Energy Technology, Chittagong University of Engineering & Technology, Chattogram-4349, Bangladesh.

ABSTRACT

A reliable rock brittleness index (BI) profile plays a key role in maintaining safe drilling operation and hydraulic fracturing during hydrocarbon exploration and field development. Also, BI is vital for mining projects such as tunneling and rock mass excavation for mineral extraction. Instead of destructive testing measurements for BI, the machine learning (ML) algorithms are widely used to obtain continuous BI profile using real field logging data. The major objectives of this study are a) to evaluate the efficacy of different ML algorithms and b) to find the most contributing geophysical logging parameters for predicting BI of clastic sedimentary formation. The data-driven computational models are developed and optimized using ML techniques of extreme gradient boosting (XGB) and random forest (RF) algorithms along with support vector regression (SVR) using a list of processed 1550 datasets from logging parameters. The explainable AI (XAI) of Shapley Additive Explanations (SHAP) is applied to identify features of importance to ensure transparency and interpretability of the BI model's predictions. Results revealed that XGB outperforms with high determination of coefficient (99%) and least statistical error compared to RF and SVR algorithms to predict BI using logging data of offshore Volve oil field. Based on XGB-based SHAP analysis, it is found that acoustic shear and compressional travel times are most contributing variables for BI forecasting other than deep resistivity, gamma ray, bulk density and neutron porosity. These features importance also verified and confirmed by SHAP value with RF. The applied methodological strategies and research findings from the study could be applied to refine the data-driven model architectures for better performance and reliability in forecasting geomechanical properties and formation evaluation with precise subsurface datasets in petroleum and mining operations.

Keywords: Logging Data, Gradient Boosting, Random Forest, Artificial Intelligence, Geomechanics.

1. Introduction

1.1. Background of Study

The rock brittleness index (BI) is a critical geomechanical property that influences hydraulic fracturing efficiency, fracture propagation, safe drilling operation and reservoir stimulation strategies, particularly in unconventional tight sands and shale formation plays. The ductile or brittle behavior of the zone of interest for hydraulic fracturing can be distinguished by BI. In enhanced oil recovery (EOR) from unconventional reservoir the hydraulic fracturing efficiency depends on the suitable selection of sweet-spot-zones for fracture initiation and perforation is based on rock BI [1]. When the static measurement of BI is unattainable due to the lack of subsurface precise rock sample collection along with standard core specimens, the formation logging variables are the alternative option to obtain BI. In recent years, machine learning (ML) techniques have been applied widely for analyzing petroleum and mining problems, including prediction of rock strength parameters and BI [1-2, 4-5]. By integrating geomechanical and logging parameters with ML approaches, an improved and reliable predictive model can be developed to obtain BI. In developing predictive models for geomechanical rock strength properties, feature selection and variable ranking play a vital role by systematically identifying and removing predictor variables that contribute the least, thereby enhancing model accuracy, reducing overfitting, and improving interpretability [4-5].

1.2 Literature Review and Problem Identification

The prediction of brittleness from well logs has been accomplished through the use of artificial intelligence (AI) techniques such as artificial neural network (ANN), support vector regression (SVR), least square support vector machine (LSSVM), random forest (RF), fuzzy logic (FL), adaptive neuro-fuzzy interference system (ANFIS) and extreme gradient boosting (XGB) of which they have seen substantial expansion in their applications within the field of geomechanics in petroleum and mining engineering [5]. In this paper, some selective scholars outcomes and limitations are explained [6-13] while predicted BI using core and log data such as gamma-ray (GR), deep resistivity (ILD), bulk density (DEN), dry density (DD), neutron porosity (NPHI), absolute porosity (PHI), acoustic compressional travel time (DTC) or velocity (Vp) as well as shear travel time (DTS) or velocity (Vs). Ghobadi and Naseri [6] adopted multiple regression and ANN techniques to assess the BI of porous limestone using outcrops geomechanical properties, including compressive strength, tensile strength, Vp, DD, PHI. Regression-based results, they concluded that integrating three input variables of PHI, Vp and DD performed better performance with coefficient of determination (R2) values of 69.9% for freeze-thaw cycle testing to measure BI. A simplified ANN model is built by Kunda and Absury [7] to predict BI using rock density, elastic properties, Vp and Vs. They concluded that the ANN-based model is performed better for BI than conventional and multiple regression techniques. Ye et. al. [8] developed computational models for predicting BI by combining back propagation-ANN with principal component analysis (PCA) using petrophysical well logs. It is concluded that the hybrid model of ANN with PCA is more accurate. Later, Wood [2] established a correlation free supervised basis data-matching model using transparent open-box algorithm to predict BI from logging and core data from lower formation of Barnett shale. Author claimed that five basic well logs of GR, ILD, DEN, NPHI, DTC are essential to find accurate models of BI. Also, Mustafa et. al. [9] investigated the data-driven computational models for BI prediction by adopting ANN and ANFIS techniques with petrophysical logs from three wells. It is revealed that ANN is better predictor than ANFIS with results of RMSE, AAPE and R^2 values of 3.78, 1.98, 94.5% for ANN and 3.51, 1.81, 95.1% for ANFIS. Hassan et. al. [10] utilized ANN, SVM, FL techniques to build data-driven model based on rock elemental composition to predict the continuous profile of calcite, clay and quartz minerals. After that they employ the predicted elemental profiles, the BI is estimated without determining the mineralogy of the rock composition. On top of that, the authors advanced another model to examine the BI profile accuracy to capture mineralogical behavior which showed high R^2 of 0.90. Zamanzadeh et. al. [1] constructed hybrid models using ML algorithms to predict BI for selecting drilling bits and hydraulic fracturing depth interval. Based on the simulated results, they claimed that five logging features of depth, Vp, Vs, DEN, GR are the most contributing to predict BI using hybrid model of LSSVM with cuckoo optimization algorithm. Ore and Gao [11] constructed ML-based predictive models to obtain BI for Marcellus shale using logging parameters of DTC, GR DEN, GR and NPHI. It is found that gradient boosting performed excellently compared to SVR and ANN. It is claimed that DTC is the significant log variable with mutual information concept. Fereidooni and Karimi [12] adopted SVR, FIS and regression tree models using 39 datasets from igneous, metamorphic and sedimentary rocks. They concluded that SVR outperformed compared to models to predict BI for complex lithology. Recently, Wu et al [13] investigated the effects of sedimentary sample texture on BI using ANN. It is revealed that limestone rock samples with mudstone texture have the slightest impact on BI. Above all studies did not attempt to apply explainable AI (XAI) like Shapley Additive Explanations (SHAP) score analysis to find feature importance, how they contributed for enhancing BI estimation. To the best of authors knowledge, still there is a research scope in the context of log data-driven computational models and features selection with explainable AI of SHAP with XGB and RF for enhancing brittleness index profile for clastic sedimentary rocks have not been examined broadly.

1.3 Research Objectives

This study goal is to deliver a continuous formation brittleness index (BI) estimation with addressing of significant and least features (logging parameters) in a time- and money-efficient manner. The major objectives of this study are listed below:

- to develop data-driven computational models and evaluate their efficacy with different ML algorithms,
- to find the most contributing logging parameters for predicting BI of clastic sedimentary rocks.

2. Materials and Methods

2.1 Studied Location and Data Acquisition

The geophysical well logs are collected from the offshore Volve oil field, which is located in the Norwegian North Sea, 200 km west of Stavanger, Norway. The stratigraphic section measured depth interval from 11,250 ft to 11,948 ft is characterized mainly by sandstone deposits with subordinate claystone interbeds [14-15]. In the study, the available geophysical well logs are gamma-ray (GR, API), bulk density (DEN, g/cc), neutron porosity (NPHI, %), deep resistivity (ILD, Ω-m) acoustic compressional travel time (DTC, µs/ft) and shear travel time (DTS, µs/ft). It is noteworthy that this section was drilled with 8-inch bit size using oil-based drilling fluid. Notably, when both Young's modulus (E, GPa) and Poisson's ratio (ν, fraction) decrease, the rock is expected to exhibit greater brittleness. Among these parameters, E plays a more significant role than ν in predicting brittleness. After careful logging data screening and ensuring data quality, using formation bulk density and acoustic travel times, the dynamic E and ν of studied sedimentary rocks are determined through the respective equations provided below:

$$E_d = \frac{DEN}{(DTS)^2} * \left(\frac{3(DTS)^2 - 4(DTC)^2}{(DTS)^2 - (DTC)^2}\right) \quad (1)$$

$$\nu_d = \frac{0.5(\frac{DTS}{DTC})^2 - 1}{(\frac{DTS}{DTC})^2 - 1} \quad (2)$$

Furthermore, the dynamic rock brittleness index (BI, fraction) is obtained from the normalized E and ν values by following equation [1]:

$$BI = \frac{E_d + \nu_d}{2} \quad (3)$$

2.2 Data-driven Model Development

The data quality is ensured before data-driven model development using bagging and boosting-based machine learning (ML) techniques of random forest (RF) and extreme gradient boosting (XGB) algorithms along with support vector regression approach. After removing noise and outlier datasets, a list of 1550 datasets adopted for each feature (x_i) and classification into training and testing phases while predicting BI. To implement the data-driven model by integration of ML and logging data, the major working steps are:

i) data collection,
ii) data preprocessing and quality assurance,
iii) data stratification and train the model,
iv) model optimization with tunned parameters,
v) model evaluation, and
vi) validate the model outcomes.

The theoretical explanation and model constraint of SVR, RF and XGB algorithms could be accessed from the available literature [11, 13,15]. The simplified mathematical expression of SVM, the following function in Equation (4) can be written, when $x_i \in R^n$ for input vectors, $\varphi(x_i)$ for feature mapping, ω for weight factor, b for bias term, and $BI_i \in R$ for output:

$$BI(x_i) = \omega^T \varphi(x_i) + b; \quad (4)$$

Considering the average value of the output from all the trees, the final equation of RF can be expressed in Equation (5) to obtain the output (BI) from the RF with a regression tree map of the x input variable ($T_m(x)$) for a number of features (m) and number of tress (M):

$$BI = \frac{1}{M}\sum_{i=1}^{M} T_m(x) \quad (5)$$

Similarly, when $f_m(x_i)$ is the prediction from the k number of regression trees, and space of regression trees is F, the global form for XGB can be written in Equation (6) for output, BI:

$$BI = \sum_{k=1}^{M} f_k(x_i);\ f_k \in \mathrm{F} \quad (6)$$

All data-driven models are trained in uniform input features and fine-tuned via hyperparameter optimization to reduce prediction error and improve generalization. The end-to-end modeling style is designed not only to compare predictive accuracy between algorithms but also to expose the underlying petrophysical drivers of brittleness. Through the combination of high-fidelity prediction and explainable feature attribution, the technique supports operational decision-making along with scientific insight into formation mechanics. The different sets of hyperparameters for each data-driven predictive model is fine-tuned then optimized models are adopted for testing the model reliability and robustness. Furthermore, an effective predictive model demonstrates strong performance through a high coefficient of determination, R^2 (close 1) and minimal error metrics, namely root mean square error (RMSE), average absolute relative percentage error (AAE), and maximum relative percentage error (MAPE). To further validate model performance beyond conventional error metrics, customized Taylor diagram [17] is used to graphically assess the statistical correlation of predicted BI values with measured test data. The diagram presents standard deviation (STD), correlation coefficient, and RMSE in one format, offering a compact assessment of model precision.

2.3 Parametric Sensitivity Analysis for Feature Selection using XAI

For assessing interdependencies among geophysical features (parameters) and how they influence BI, Pearson's correlation coefficient (CC) is analyzed for the entire dataset to find their relationship. Notably, the explainable artificial intelligence (XAI) technique of SHAP with machine learning algorithms utilized to find the features importance while predicting BI of clastic sedimentary rocks. The SHAP value is evaluated by following equation for model output of $f(P)$ is mentioned below [18]:

$$\mathrm{SHAP}_i = \sum_{P\subseteq M\setminus i} \frac{|P|!\,(M-|P|-1)!}{M!}\left[f(P\cup i) - f(P)\right] \quad (7)$$

where, P signifies a set of non-zero directories and M denotes the set of all input logging features.

Moreover, the bar plots are illustrated to show the average absolute SHAP values for each predictor feature, which measures how much each feature contributes to the model's predictions. A higher value signifies a stronger impact of predictor variable on BI prediction, and vice versa. Additionally, the beeswarm diagrams provide a visualization of the SHAP values for each feature across all testing instances. Each point represents one instance, positioned along the x-axis according to its SHAP value for the corresponding feature. The point color encodes the feature's magnitude, where warmer tones (red) represent higher values and cooler tones (blue) represent lower values.

3. Results and Discussion

3.1 Data Distribution

A list of datasets of 1,550 are collected from high-resolution well log measurements to adopt six predictor variables such as compressional travel time (DTC, µs/ft), shear travel time (DTS, µs/ft), gamma ray (GR, API), bulk density (DEN, g/cc), neutron porosity (PHIN, %), and deep resistivity (ILD, Ω.m) to obtain output variable brittleness index (BI). The summarized result for descriptive statistical parameters is listed in Table 1 to measure the central tendency and variability focusing on spreading out the datasets.

Table 1 Statistical distribution of considered variables.

Variable	Min	Max	Mean	Median	STD
DTC (µs/ft)	59.42	89.44	72.59	72.42	5.45
DTS (µs/ft)	96.90	156.82	124.41	124.01	10.92
GR (API)	24.45	127.05	55.73	46.75	24.39
DEN (g/cc)	2.313	2.68	2.53	2.53	0.09
PHIN (%)	5.48	24.38	14.49	14.69	4.00
ILD (Ω-m)	0.59	3.79	1.79	1.53	0.84
BI (frac.)	0.04	0.77	0.37	0.35	0.13

These features are selected based on their physical relevance to rock elasticity, lithological effects, porosity, and fluid content to influence brittleness behavior of studied formation depth interval of 11433 ft to 11948 ft. The scaled BI ranges from 0.04 to 077 in the oil field well which is determined through an elastic modulus (E) and Poison's ratio (ν)-based equation to facilitate ML-based regression modeling and interpretability. Notably, the range of E and ν values are 22.41-62.12 GPa and 0.10-0.31, respectively.

2.2. Effectiveness of Data-driven Predictive Models

To accomplish intensive training and testing of the model through trial-and-error approach, the datasets is split into training and testing subsets, like 65% (1,007 samples) to train the ML-based models and 35% (543 samples) to test and validate the models. The fairly large test set is chosen to achieve statistically significant performance metrics and prevent the risk of overfitting. The optimized hyperparameters for radial basis kernel function (RBF) with SVR model are regularization factor (C), cache_size, degree, epsilon (ϵ) and tolerance of 1.0, 200, 3, 0.1 and 0.001, respectively. The tuned parameters of RF model are number of trees of the forest (n_estimators), max_features, min_samples_leaf and split of 100, 1.0, 1 and 2, respectively. The performance of ML-based predictive models is listed in Table 2 for testing phases. Out of the three ML-models in the study, XGB performed best on all metrics with least error of RMSE (0.01), MAPE (2.23%), AAPE (2.19%), and high R^2 (99.35%) compared to RF and SVR models.

Table 2 Effectiveness of ML models to predict BI.

Algorithm type	MAPE (%)	AAPE (%)	R^2 (%)
XGB	2.23	2.19	99
RF	2.45	2.44	98
SVR	9.00	8.65	92

The Taylor diagram is illustrated in Fig. 1 to show the distribution of observed BI, a standard deviation of 0.1274 is used as a baseline reference.

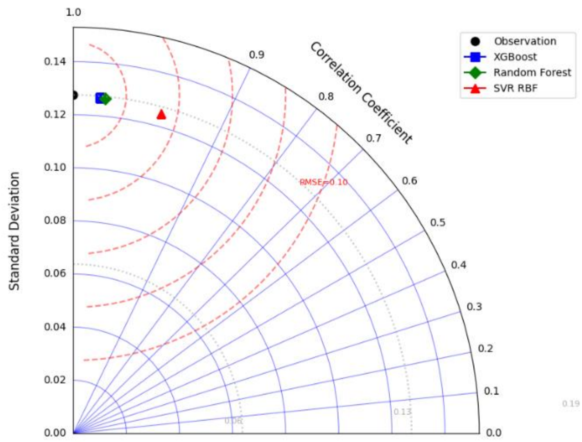

Fig. 1 An illustration of Taylor diagram to show the model performance with ML algorithms.

The XGB model is nearly as close to this with a predicted standard deviation of 0.1268 and correlation coefficient of 0.9968, indicating nearly perfect agreement in amplitude and pattern. RF also exhibited close agreement with the observed data too while SVR had slightly poorer standard deviation, 0.1249 and correlation 0.9642, although still within acceptable limits but repeated its relatively poorer performance compared to other models. The proximity of XGB and RF to the point of reference on the graph, along with their placement within the lowest RMSE contours, attest to their robustness and tractability for BI prediction. These statistics indicate that nearly perfect predictive precision and minimal variation from actual BI values with XGB algorithms. The findings highlight the significance of algorithm selection in petrophysical modeling and validate the use of gradient-boosted frameworks for geomechanical property estimation.

2.3. Effects of Geophysical Logging Parameters for BI

Pearson's correlation coefficient (CC) between features is summarized in Table 3 to find their correlation either positively or negatively. Results revealed DTC and DTS features are most significant parameters and strong negative correlated with BI with CC of 95 and 82, respectively. Moreover, GR variable is slightly negative correlated with BI with CC of 0.16.

Table 3 Tabulated form of correlation between features.

Feature	DTC	DTS	GR	DEN	PHIN	ILD	BI
DTC	1						
DTS	0.95	1					
GR	0.05	0.12	1				
DEN	-0.71	-0.64	0.38	1			
PHIN	0.73	0.73	0.35	-0.55	1		
ILD	-0.63	-0.60	0.44	0.83	-0.45	1	
BI	-0.82	-0.95	-0.16	0.57	-0.68	0.57	1

In addition, SHAP value-based features contribution to predict BI is illustrated in Fig. 2 and Fig. 3 for XGB and RF algorithms, respectively. The feature importance with SHAP profiles is suggesting consistent patterns in models with dramatic variations in how each algorithm distributes importance to geophysical logging inputs.

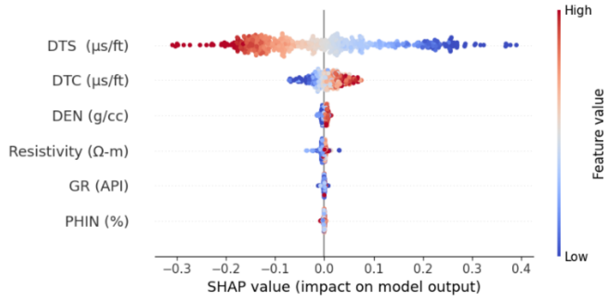

Fig. 2 An illustration of beeswarm plots with SHAP value for features evaluation using XGB algorithm.

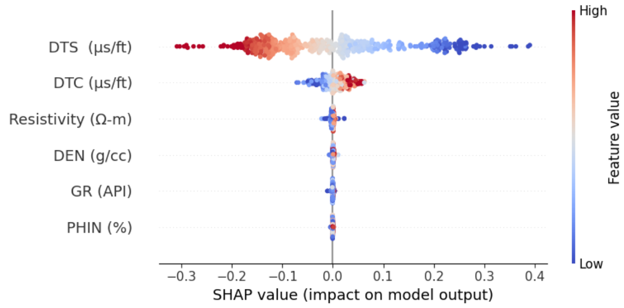

Fig. 3 An illustration of beeswarm plots with SHAP value for features evaluation using RF algorithm.

The feature of acoustic shear travel time is the highest contributor in the XGB model with a SHAP importance of 0.1150. This is a reflection of the model's reliance on elastic parameters like the shear modulus, which influences rock brittleness directly. Compressional travel time (DTC) is the second significant variable with a moderate SHAP importance of 0.0196, confirming its role in identifying complementary elastic behavior. The remaining features of DEN, deep resistivity, GR and NPHI are lest significant to obtain BI, respectively.

In addition, RF model revealed a nearly identical feature ranking, the highest-order predictors again being DTS (0.1150) and DTC (0.0196) and least significant feature is

PHIN (less than 0.003). Deep resistivity (ILD), DEN and GR are intermediate contributing features for BI estimation. The close resemblance of XGB and RF suggests that tree-based models are perfectly matched for extracting the elastic controls on rock brittleness, and deprioritizing features with lesser or indirect control. Collectively, these results affirm that elastic attributes are the predominant cause of brittleness index in this dataset while DTS and DTC are primary performing indicators out of all the explainable models. It is noteworthy that lithology and fluid-related characteristics are contributed negligibly to brittleness models, especially to XGB and RF, does show evidence for the mechanical behavior of brittleness and advocates for the important role played by shear wave data in geomechanical modeling. The ability of tree-based models to generalize across diverse petrophysical conditions further enhances their applicability in field-scale reservoir evaluation, offering a scalable and interpretable framework for brittleness screening in unconventional plays.

4. Study Limitations and Future Research Scopes

While the brittleness index (BI) prediction framework developed in this study demonstrates high accuracy and interpretability, several limitations merit consideration. The input dataset is restricted to conventional geophysical well logs, excluding mineralogical, geochemical, and laboratory-derived mechanical data that could enhance model fidelity and geological resolution. As a result, the models may not fully capture lithology-specific brittleness behaviors, particularly in formations with complex mineralogy, organic content, or anisotropic formation stress regimes. Additionally, the BI is computed using a single elastic modulus-based formulation, which, while widely accepted, may not generalize across all geological settings or account for alternative brittleness definitions based on mineral composition or empirical geophysics relationships. Notably, considering the most significant contributing predictor variables to estimate BI, the improved correlation did not perform in the current study. This study can be attempted using real field logging data with gene expression programming approach.

Additionally, deploying a broader suite of machine learning algorithms, including gradient boosting variants, and hybrid models with metaheuristic optimization, could further optimize predictive performance and robustness of the model. Finally, coupling BI prediction with facies classification, fracture modeling, or reservoir simulation workflows would extend its utility from diagnostic analysis to operational decision support in unconventional shale reservoir development.

5. Conclusions

In the study, data-driven computational models are developed and evaluated by integration of machine learning (ML) and geophysical logging data for brittleness index (BI) prediction. Based on the research findings, the following outcomes can be concluded as below:

- The extreme gradient boosting (XGB) is the more effective and generalized algorithm, and it performed with high coefficient of correlation (99.35%) and least error compared to the random forest (RF) and support vector regression (SVR).
- Taylor diagram comparison further enhanced model validity, with XGB and RF performing just a whisker away from reproducing the actual BI distribution and still achieving high correlation coefficient of 0.9968 and 0.9955, respectively compared to SVR which showed poorer generalization.
- SHAP analysis consistently identified acoustic shear travel time (DTS) as the most dominant feature in all explainable models, with SHAP values of 0.114, confirming mechanical sensitivity of BI to elastic properties.
- Compressional travel time (DTC) contributed moderately, while lithological and fluid indicators such as gamma-ray, neutron porosity, deep resistivity, and bulk density contributed negligibly, particularly in tree-based models.
- Improved BI prediction using DTS and DTC logging features with deep learning is anticipated to optimize hydraulic fracturing, reduce costs, and strengthen the economic sustainability of shale gas and sandstone reservoir development.

6. Acknowledgement

This work was partially funded by the Natural Sciences and Engineering Research Council of Canada.

7. References

[1] Zamanzadeh Talkhouncheh, M., Davoodi, S., Larki, B., Mehrad, M., Rashidi, S. and Vasfi, M., 2023. A new approach to mechanical brittleness index modeling based on conventional well logs using hybrid algorithms. *Earth Science Informatics*, *16*(4), pp.3387-3416.

[2] Wood, D.A., 2021. Brittleness index predictions from Lower Barnett Shale well-log data applying an optimized data matching algorithm at various sampling densities. *Geoscience Frontiers*, *12*(6), p.101087.

[3] Rana, M.R., Miah, M.I. and Abir, M.A.N., 2025. Experimental Study on Rock Strength Parameters for Brittleness Index Evaluation: A Case Study of the Sitakunda Anticline Area, Bengal Basin. *Journal of Petroleum and Mining Engineering*, *26*(2), pp.87-97.

[4] Miah, M.I., Ahmed, S., Zendehboudi, S. and Butt, S., 2020. Machine learning approach to model rock strength: prediction and variable selection with aid of log data. *Rock Mechanics and Rock Engineering*, *53*(10), pp.4691-4715.

[5] Miah, M.I., 2020. Predictive models and feature ranking in reservoir geomechanics: A critical review and research guidelines. *Journal of Natural Gas Science and Engineering*, *82*, p.103493.

[6] Ghobadi, M.H. and Naseri, F., 2016. Rock brittleness prediction using geomechanical properties of Hamekasi limestone: regression and artificial neural networks analysis. *Geopersia*, *6*(1), pp.19-33.

[7] Kaunda, R.B. and Asbury, B., 2016. Prediction of rock brittleness using nondestructive methods for hard rock tunneling. *Journal of Rock Mechanics and Geotechnical Engineering*, *8*(4), pp.533-540.

[8] Ye, Y., Tang, S., Xi, Z., Jiang, D. and Duan, Y., 2022. A new method to predict brittleness index for shale gas reservoirs: Insights from well logging data. *Journal of Petroleum Science and Engineering*, *208*, p.109431.

[9] Mustafa, A., Tariq, Z., Mahmoud, M.A. and Abdulraheem, A., 2022, June. Artificial intelligence approach for predicting the shale brittleness index-a middle east basin case study. In *SPE Europec featured at EAGE Conference and Exhibition?* (p. D042S011R002). SPE.

[10] Hassan, A., Chan, S., Mahmoud, M., Aljawad, M.S., Humphrey, J. and Abdulraheem, A., 2022. Artificial intelligence-based model of mineralogical brittleness index based on rock elemental compositions. *Arabian Journal for Science and Engineering*, *47*(9), pp.11745-11761.

[11] Ore, T. and Gao, D., 2023. Prediction of reservoir brittleness from geophysical logs using machine learning algorithms. *Computers & Geosciences*, *171*, p.105266.

[12] Fereidooni, D. and Karimi, Z., 2023. Predicting rock brittleness indices from simple laboratory test results using some machine learning methods. *Geomechanics and Engineering*, *34*(6), pp.697-726.

[13] Wu, H., Wu, D. and Wang, L., 2025. Assessment and Estimation of Sedimentary Rocks' Brittleness Using Machine Learning and Statistical Methods. *Mining, Metallurgy & Exploration*, pp.1-31.

[14] Oloruntobi, O., Onalo, D., Adedigba, S., James, L., Chunduru, R. and Butt, S., 2019. Data-driven shear wave velocity prediction model for siliciclastic rocks. *Journal of Petroleum Science and Engineering*, *183*, p.106293.

[15] Miah, M.I., Ahmed, S. and Zendehboudi, S., 2021. Model development for shear sonic velocity using geophysical log data: Sensitivity analysis and statistical assessment. *Journal of Natural Gas Science and Engineering*, *88*, p.103778.

[16] Miah, M.I., Abir, M.A.N. and Shuvo, M.A.I., 2025. Improved model development and feature ranking for rock permeability prediction by coupling petrophysical log data and ensemble machine learning techniques. *Earth Science Informatics*, *18*(1), p.93.

[17] Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. *Journal of geophysical research: atmospheres*, *106*(D7), pp.7183-7192.

[18] Rahaman, M.S. and Miah, M.I., 2025. Data-driven resilient model development and feature selection for rock compressive strength prediction using machine learning and transformer techniques. *Earth Science Informatics*, *18*(3), p.293.