Investigating the Applications of Machine Learning Techniques to Predict the Rock Brittleness Index

Deliang Sun [1], Mahshid Lonbani [2], Behnam Askarian [3], Danial Jahed Armaghani [4],*[iD], Reza Tarinejad [5][iD], Binh Thai Pham [6][iD] and Van Van Huynh [4]

[1] The Key Laboratory of GIS Application Research, Chongqing Normal University, Chongqing 401331, China; sundeliang2002@aliyun.com

[2] Faculty of Management, Universiti Teknologi Malaysia (UTM), Johor 81310, Malaysia; Lonbanimahshid@gmail.com

[3] Department of Electrical and Computer Engineering, Texas Tech University, Lubbock, TX 79409, USA; Behnam.askarian@ttu.edu

[4] Modeling Evolutionary Algorithms Simulation and Artificial Intelligence, Faculty of Electrical & Electronics Engineering, Ton Duc Thang University, Ho Chi Minh City 758307, Vietnam; huynhvanvan@tdtu.edu.vn

[5] Department of Civil Engineering, University of Tabriz, 29 Bahman Blvd, 51666 Tabriz, Iran; r_tarinejad@tabrizu.ac.ir

[6] Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam; phamthaibinh2@duytan.edu.vn

* Correspondence: danialjahedarmaghani@tdtu.edu.vn

check for updates

Abstract: Despite the vast usage of machine learning techniques to solve engineering problems, a very limited number of studies on the rock brittleness index (*BI*) have used these techniques to analyze issues in this field. The present study developed five well-known machine learning techniques and compared their performance to predict the brittleness index of the rock samples. The comparison of the models' performance was conducted through a ranking system. These techniques included Chi-square automatic interaction detector (CHAID), random forest (RF), support vector machine (SVM), K-nearest neighbors (KNN), and artificial neural network (ANN). This study used a dataset from a water transfer tunneling project in Malaysia. Results of simple rock index tests i.e., Schmidt hammer, p-wave velocity, point load, and density were considered as model inputs. The results of this study indicated that while the RF model had the best performance for training (ranking = 25), the ANN outperformed other models for testing (ranking = 22). However, the KNN model achieved the highest cumulative ranking, which was 37. The KNN model showed desirable stability for both training and testing. However, the results of validation stage indicated that RF model with coefficient of determination (R^2) of 0.971 provides higher performance capacity for prediction of the rock *BI* compared to KNN model with R^2 of 0.807 and ANN model with R^2 of 0.860. The results of this study suggest a practical use of the machine learning models in solving problems related to rock mechanics specially rock brittleness index.

Keywords: rock brittleness index; machine learning techniques; artificial neural networks; random forest; K-nearest neighbors; support vector machine

1. Introduction

In underground space and excavation related projects, brittleness of the rock is considered as one of the most important properties of the rock mass. Having an appropriate insight on rock brittleness in other fields of engineering also help engineers alleviate the issues related to brittleness. For example, the acquisition of sufficient knowledge on the rock brittleness by oil and gas engineers could help

them to evaluate the wellbore stability as well as appraise the performance of a hydraulic fracturing job [1]. Moreover, the brittleness regulates the properties of the shale rocks mechanic. At the same time, by employing several parameters such as the carbonates, volumetric fraction of strong minerals, weak elements and pores, Young modulus and strength of these properties can be defined [2]. In deep underground engineering, brittleness is a critical factor to assess the stability of the surrounding rock mass [3].

Besides, many disasters related to rock mechanics like rock-bursts may stem from brittleness [4–6]. Several studies showed that brittleness is also an important factor to estimate the tunnel boring machine (*TBM*) and road-harder cutting performance [7]. In addition, it defines the excavation efficiency of drilling, which considerably influences coal mining [8]. Hence, the assessment of rock brittleness is necessary for geotechnical and rock mechanics projects [5]. However, despite the fact that the brittleness is an important parameter for designing civil and mining engineering projects, but according to Altindag [9], there is still no consensus on definition and measurement standards of this phenomenon. Hence, according to Yagiz [10], various rock properties influence rock brittleness. Some studies have related brittleness to the lack of ductility or ductility inversion [11]. Ramsey [12] also defined the brittleness as breaking of inter-particle cohesion of a rock. In addition, Obert and Duvall [13] pointed out that brittleness is the inclination of a material, such as cast iron or many types of rocks, to be split following pressure equivalent or higher than the material yield stress. A highly brittle rock typically has the following features: (a) failure without a considerable force, (b) generation of small particles, (c) great ratio of compressive to tensile strength, (d) great firmness, (e) great interior friction angle, and (f) production of fully developed fractures following hardness lab experiments [14]. It seems that the majority of studies that were conducted on rock brittleness index (*BI*) were based on relationship between tensile and uniaxial compressive strengths of the rock samples [15–17]. However, a few studies have presented and suggested a relationship between *BI* with other rock properties such elasticity modulus, hardness, Poisson's ratio, internal friction angle, and quartz content [18]. The performance of these models have reported not capable enough to predict the *BI*. This is because of the fact that most of them used one or two dependent parameters [8,17]. It seems that the use of multi-inputs predictive systems in estimating *BI* of the rock would be great to receive a higher degree of accuracy compared to simple regression models.

Recently a large number of studies used soft computing (SC), machine learning (ML) and artificial intelligence (AI) techniques to solve problems related to science and engineering fields [19–39]. However, a limited number of studies relevant to these methods have been conducted to predict the rock *BI* in literature. Kaunda and Asbury [40] employed an artificial neural network (ANN) technique to predict the rock *BI* using system inputs such as the velocity of the S and P waves, Poisson's ratio, elastic modulus and unit weight. Yagiz and Gokceoglu [8] estimated the rock *BI* by constructing fuzzy inference system (*FIS*) and non-linear regression analysis. The inputs that were used to develop these models were the unit weight, uniaxial compressive strength (*UCS*) and Brazilian tensile strength (*BTS*) of the rock. They concluded that the *FIS* model is an applicable technique in order to be used in the same field for further studies. Koopialipoor et al. [16] proposed predictive equations for calculation of rock *BI* as a function of intact rock properties including rock density, Schmidt hammer rebound number and p-wave velocity. They used a hybrid approach and combined the firefly algorithm and ANN models to develop the equitation. Khandelwal et al. [17] examined the feasibility of genetic programming model for predicting the brittleness of intact rocks. Their study used multiple input variables including *UCS*, *BTS* and unit weight to forecast the *BI* of the rock mass.

While several previous studies acknowledged the suitability of ML techniques for solving the engineering problems, several ML techniques remained unused or barely applied to predict the rock *BI*. To the authors best of knowledge, no study is available which examined the feasibility of well-known ML techniques such as chi-square automatic interaction detector (CHAID), random forest (RF), support vector machine (SVM), and K-nearest neighbors (KNN) for predicting the *BI*. Thus, in this study the abovementioned ML techniques plus ANN technique (as a benchmark one in field of ML) were

employed for *BI* prediction purpose. The performance of each model was evaluated through five performance indices and a gain chart. Additionally, three best models of this study are discussed in more details.

2. Methodology

2.1. Models Developed

The models that were developed in this study are CHAID, RF, SVM, KNN, and ANN. The CHAID is from decision tree family which produce non-binary tree structure. This technique that was developed by Kass [41] employs a chi-square test to produce multiple sequential combinations and splits, and finally a single decision tree. Typically, the decision tree techniques are susceptible to overfitting. However, the CHAID automatically prunes the tree to alleviate the overfitting phenomenon. Moreover, the CHAID generates a number of rule sets and each of these rule sets has a confidence level and accuracy.

While the single-based decision trees are easy to implement and understand, these techniques are prone to result in different generalization behaviour with small changes to the training data [42]. Indeed, these techniques are viewed as unstable and high variance. The RF technique is extensively efficient to remedy the abovementioned shortcomings of single-based decision trees. This technique was developed by Breiman [43] and is an ensemble-based approach (Figure 1). The RF generates more accurate prediction results compared to single-based trees since it combines a huge number of single trees. It is worth pointing out that the RF enjoys a bagging approach to create each number of an ensemble from diverse datasets. This approach randomly chooses from the space of single decision trees and generates almost identical (low diversity) predictions.
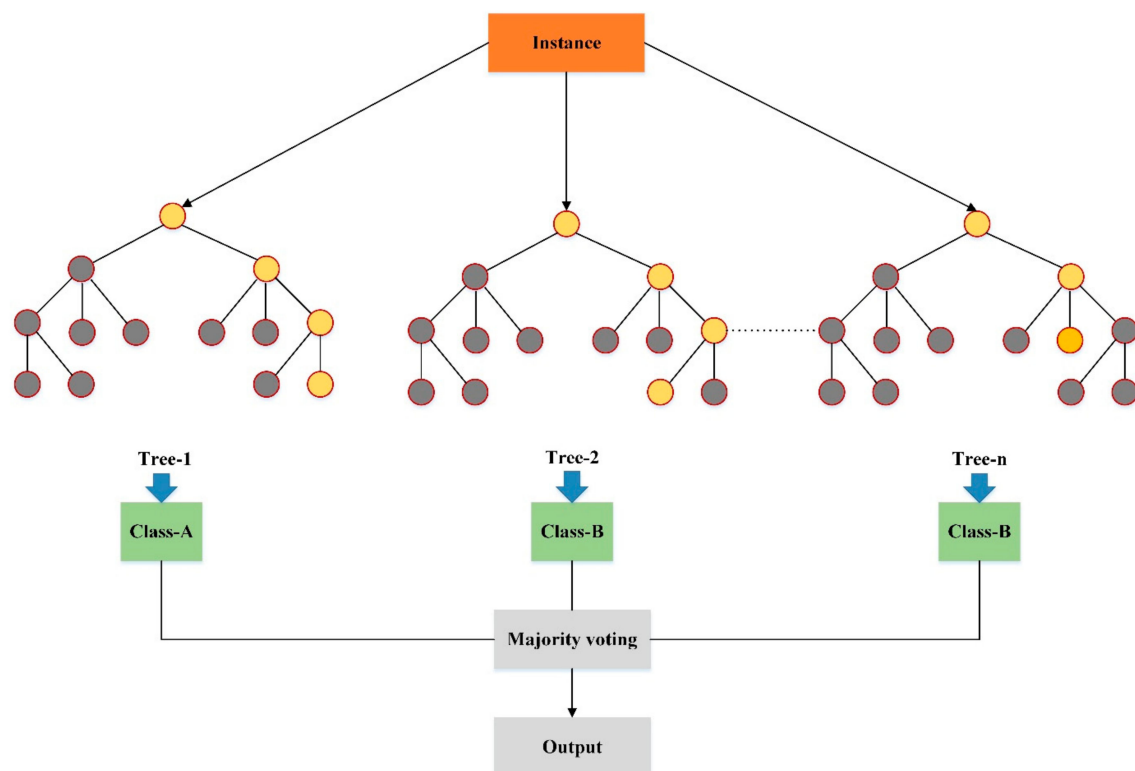

Figure 1. RF structure.

Other ML techniques such as KNN, SVM, and ANN are also powerful tools for classification and regression analysis in civil and mining problems [35]. KNN is an easy to implement, simple and effective data mining algorithm [44]. The basic theory behind the KNN is discovering a group of "*k*"

samples (e.g., employing the distance functions) which has the nearest distance from unknown samples in the calibration dataset. Moreover, the KNN identifies the class of unknown samples among the "*k*" samples by calculating the average of the response variables [45]. Thus, the "*k*" plays an important role in the KNN performance [46].

Concerning the SVM, this technique is capable to handle the high dimensional and linearly non-separable datasets [47,48]. In addition, it can reduce the error for training and testing datasets, as well as the model complexity [49]. According to Cortes and Vapnik [50], statistical learning theory is the basic theory behind the SVM. In addition, the performance of SVM is influenced by Kernel functions, such as linear, radial basis function, sigmoid and polynomial [51]. It is noteworthy to mention that the SVM aims to determine a perfect separation hyper plane that can distinguish the two classes [51]. The SVM regression aims to discover the largest margin. Figure 2 shows a typical structure of SVM.
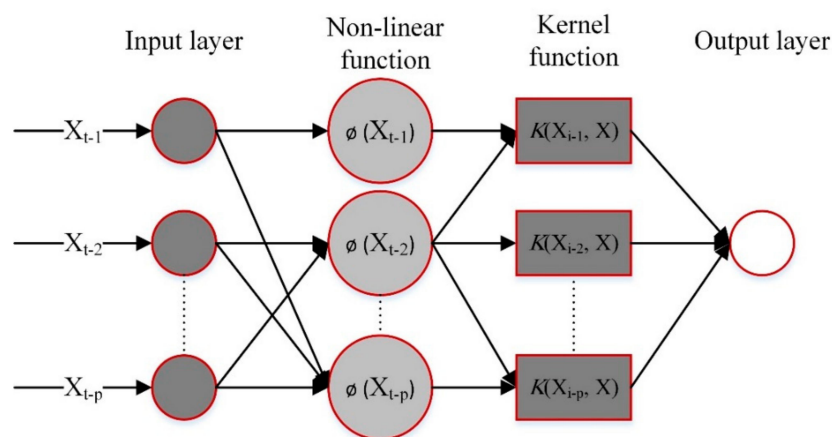

Figure 2. Typical structure of SVM.

In terms of ANN, this technique is a kind of artificial intelligence which emulates some functions of an individual's mind. Typically, the ANN is tended to sort experiential knowledge [52]. This technique includes a series of layers, and each layer includes a sequence of neurons. These neurons in every layer are connected thorough weighted links to all neurons on the previous and following layers [52–55].

A positive weight reveals an excitatory association, whereas a negative weight reveals an inhibitory association. A typical ANN includes three layers, i.e., input layer, hidden layer, and an output layer [56–60]. This structure is shown in Figure 3 for more illustration.
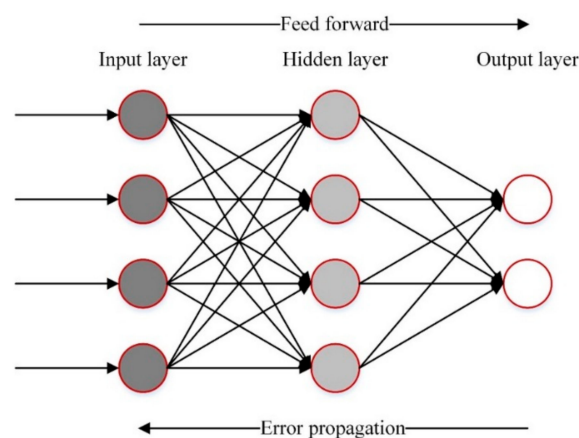

Figure 3. Typical three-layer neural network.

2.2. Data and Case Study

The data of this study was acquired from the Pahang-Selangor tunnel, Malaysia (Figure 4). The main aim of constructing this tunnel was to provide a flow path of fresh water. The tunnel specifications are as follow: (1) diameter: 5.2 m; length: 44.6 km; and longitudinal gradient: 1/1900. In addition, under free-flow conditions, the maximum allowable discharge of the tunnel is 27.6 m^3/s. In order to excavate the tunnel, three different *TBMs* were used for about 35 km of the tunnel. The remainder of the tunnel was excavated using the drilling and blasting method. The geological units include granite, metamorphic and some sedimentary rocks; though, most of the rocks excavated with the abovementioned method is comprised of granite. Many geotechnical and geological investigations were conducted in the tunnel to collect rock block samples for testing. Finally, in multiple locations of *TBMs* site, more than 100 granite block samples were obtained from the tunnel face. A robust procedure from the International Society for Rock Mechanics [61] was followed for preparing the samples to test. Several lab tests were conducted on the samples, including density (in dry condition), Schmidt hammer rebound number (R_n), uniaxial compression strength (σ_c), tensile strength (σ_t), point load index (Is_{50}), and p-wave velocity (V_p). In this study, the *BI* values were calculated according to the following equation [62]:

$$BI = \frac{\sigma_c}{\sigma_t} \tag{1}$$

where, σ_c and σ_t are the uniaxial compression strength and tensile strength, respectively.

Thus, considering and selecting *BI* values as model output, four parameters of density, Schmidt hammer rebound number, point load index and p-wave velocity were set as inputs in the form of a database with 110 datasets. The range, mean, unit and symbols of inputs and output parameters in this study are tabulated in Table 1. According to this table, average values of 5491.6 m/s, 2.59 g/cm^3, 40.5, 3.6 MPa, and 15.5 were obtained for V_p, D, R_n, Is_{50} and *BI*, respectively. In the next section, modeling procedure in approximating *BI* as a function form of f (V_p, D, R_n, and Is_{50}) and the obtained results will be presented in detail.

Table 1. The Range, Mean, Unit, Category and Symbol of Inputs and Output Parameters in Predicting *BI* of the Rock Samples.

Parameter	Symbol	Unit	Category	Min	Max	Mean
P-wave velocity	V_p	m/s	Input	2870	7702	5491.6
Density	D	g/cm^3	Input	2.37	2.79	2.59
Schmidt hammer rebound number	R_n	-	Input	20	61	40.5
Point load strength	Is_{50}	MPa	Input	0.89	7.1	3.6
Brittleness index	BI	-	Output	8.90	24.01	15.5

3. Modelling Process and Results

The present study developed five ML models to predict *BI* of the rock material. To develop the models, a database contained 110 datasets was used. These data were split into the train and test with the ratio of train to test being 70%: 30%. Thus, 77 samples for training and 33 samples were used for testing. As pointed out earlier, five ML models including, RF, CHAID, SVM, KNN and ANN were developed to estimate *BI* of the rock. Each of these models were evaluated using a simple ranking system and a gains chart. The three best models are discussed in more detail.
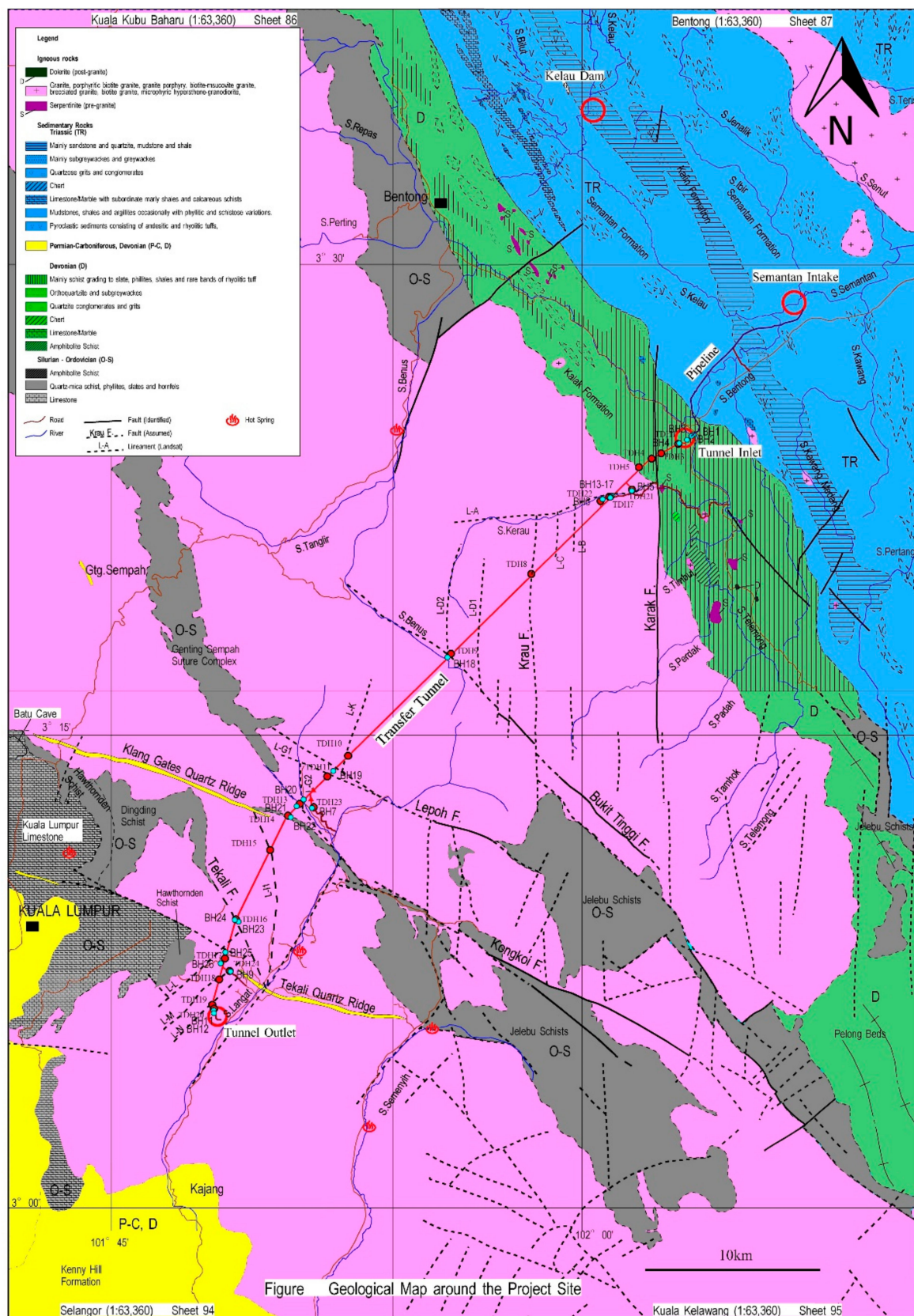
Figure 4. Geological map of tunnel location and its route.

3.1. Evaluation of the Developed Models

Once the models have been developed, the accuracy performance of each model was evaluated using five well-known indices i.e., coefficient of determination (R^2), root mean square error (*RMSE*), mean absolute error (*MAE*), variance account for (*VAF*), and *a20-index*. The formula that was used for calculating the mentioned performance indices are presented in Equations (2)–(6). This study also employed an easy to understand ranking system which ranked each model developed using the above-mentioned performance criteria for both training and testing stages. For each criteria, the ranking system first sorted the models based on their obtained values, then assigned the highest rank (5) to the best value and the lowest rank (1) to the worst value. Final rank of each model was calculated through summing the ranking values for both training and testing stages (Equation (6)):

$$R^2 = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \overline{y})^2} \tag{2}$$

$$VAF = \left[1 - \frac{var\ (y - y\prime)}{var\ (y)} \right] \times 100 \tag{3}$$

$$MAE = \frac{1}{N} \sum_{j=1}^{N} |y_i - y_j| \tag{4}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y - y\prime)^2} \tag{5}$$

$$a20 - index = \frac{m20}{N} \tag{6}$$

where y denotes the measured values, $\overline{y}$ and $y\prime$ indicate mean and predicted of the y, respectively, N denotes the total number of data, $m20$ shows number of samples with value of experimental value/predicted value between 0.8 and 1.20:

$$Final\ rank\ of\ model = \sum_{\substack{1 \leq i \leq 5}}^{\substack{1 \leq j \leq 2}} R_{ij} \tag{7}$$

where i denotes the indices, j shows the dataset, the R shows model's ranking.

The values and ranks of the performance indices and models are presented in Table 2. The results of this evaluation showed that the KNN model achieved the highest final rank (37). This model was followed by RF (34) and ANN (33), respectively. For the training dataset, the RF obtained the highest rank (25) while the SVM obtained the lowest rank (5). For the testing dataset, the ANN achieved the highest rank (22) while the CHAID achieved the lowest rank (6). Turning to the performance indices, the RF outperformed other models developed for the training dataset. However, for the testing dataset, the ANN achieved the best ranks for three indices including R^2, *RMSE*, and *a20-index*. The ANN also achieved the second-best rank for *VAF*. Based on these discussions and rank values, three models of RF, KNN, and ANN were selected to be discussed in more details in the following sections.

The authors also used a gain chart (Figure 5) to compare the performance of the models proposed for both training and testing datasets. Gains are estimated as (number of hits in quantile/total number of hits) × 100%. Here, it is necessary to mention that "hit" refers to the success of a model to predict the values greater than the midpoint of the fields range (*BI* > 16.458). In this chart, the blue line denotes the perfect model which has perfect confidence (where hits = 100% of cases), the diagonal red line denotes the at-chance model, and the other five lines in the middle represent the models developed in this study.

Table 2. Evaluation of Models Developed Using Five Performance Indices.

Performance Index	RF				CHAID				KNN				SVM				ANN			
	TR		TE		TR		TE		TR		TE		TR		TE		TR		TE	
	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank	Value	Rank
R^2	0.89	5	0.75	2	0.77	3	0.74	1	0.81	4	0.84	4	0.73	1	0.84	3	0.75	2	0.85	5
RMSE	1.08	5	1.75	2	1.39	3	1.8	1	1.33	4	1.44	3	1.57	1	1.42	4	1.50	2	1.39	5
VAF (%)	86.84	5	80.5	2	78.31	3	75.1	1	80.62	4	86.1	3	72.62	1	87.5	5	74.60	2	87	4
MAE	0.91	5	1.36	1	1.14	4	1.35	2	1.16	3	1.15	4	1.32	1	1.06	5	1.29	2	1.16	3
a20-index	1.00	5	0.91	2	0.95	2	0.84	1	0.99	4	0.97	4	0.91	1	0.96	3	0.97	3	0.97	5
Sum of the ranks	TR	25	TE	9	TR	15	TE	6	TR	19	TE	18	TR	5	TE	20	TR	11	TE	22
Final rank		34				21				37				25				33		

Perfect R^2 = 1; Perfect *RMSE* = 0; Perfect *VAF* = 100%; Perfect *MAE* = 0; *a20-index* = 1. Training dataset = TR; Testing dataset = TE
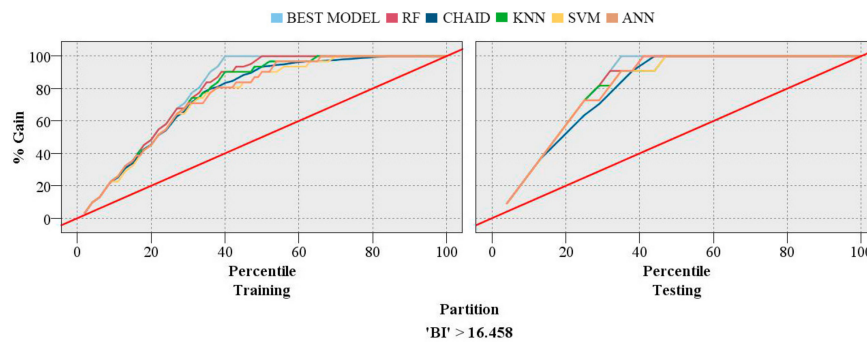
Figure 5. Evaluation of the models proposed using a gain chart.

Typically, the higher lines show better models, particularly on the left side of the chart. To compare a model developed and the at-chance model, the area between a model and the red line can be used. In fact, this area identifies how much better a proposed model is compared to the at-chance model. Additionally, the area between a model proposed and the perfect model identifies where a proposed model can be improved.

For training stage, it is shown that the perfect model has correctly identified 100% of the samples, which had the *BI* of greater than 16.458, at the percentile of 40%. The RF model was the closest follower of the perfect model and correctly identified 100% of the samples, which had the *BI* greater than 16.458, at the percentile of 50%. The weakest model was the CHAID which identified the hit at the percentile of 84%. For testing dataset, it can be seen that that the perfect model has correctly identified 100% of the samples, which had the *BI* of greater than 16.458, at the percentile of 35%. The perfect model was followed by the RF and ANN models which had correctly identified 100% of the samples, which had the *BI* of greater than 16.458, at the percentile of 41%. The KNN and ANN had the weakest performance and identified the hit at the percentile of 47%.

3.2. Conqueror Models

3.2.1. Random Forest Model

The RF model was developed using four input variables, including V_p, D, R_n, and Is_{50} to predict the rock *BI*. The present study employed several parameters to develop the RF model. After trial and error procedure, the number of models to build was set as 100, the sample size was set as 0.95, the maximum number of nodes was set as 10,000, and maximum tree depth and minimum child node size were set as 10 and 2, respectively. Predicted *BI* values by RF, along with their actual values for training and testing datasets, are displayed in Figure 6. The obtained R^2 values of 0.89 and 0.75 for train and test stages of RF model, respectively revealed a high and suitable accuracy level of train and test stages. In addition, the RF model identified the importance of the input variables (Figure 7). As can be seen, the R_n with importance of 0.37 was identified as the most important variable and followed by V_p with importance of 0.35 and Is_{50} with importance of 0.29. It is noteworthy to mention that the RF model did not consider *D* as an important factor.
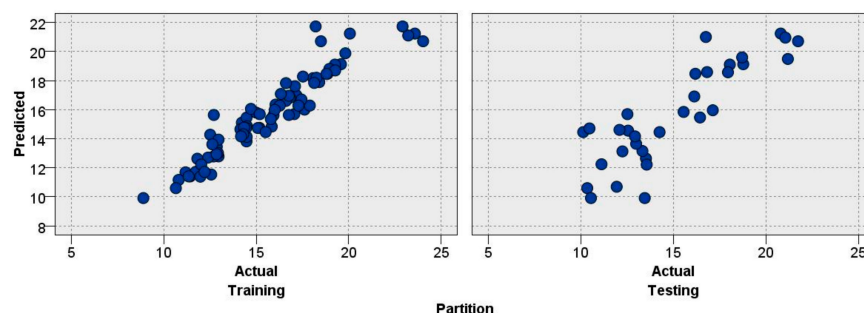

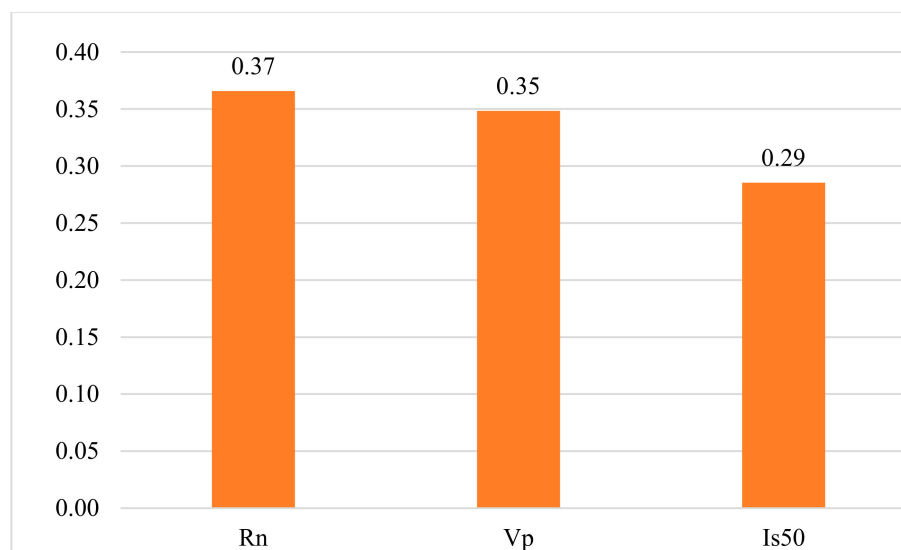**Figure 6.** Testing and training results of RF model to predict the *BI*.

Figure 7. Input variable's importance to predict the *BI* derived from the RF model.

3.2.2. ANN Model

As mentioned earlier, this study developed the ML predictive models by means of four input variables, i.e., V_p, D, R_n, and Is_{50} for predicting the rock *BI*. Here, several parameters have been used to develop the ANN model. The type of neural network model was multilayer perceptron. The study used "mean" as the default combing rule for our continuous target. Number of component models for boosting and/or bagging was set as 10. To avoid over-fitting, the over-fit prevention set was set as 30%. Different values were examined in order to determine the number of hidden neurons and in the final model, a number of 4 hidden neurons was used to predict *BI*. Figure 8 shows the suggested architecture of the ANN model with four input neurons, four hidden neurons and one output neuron in predicting *BI* of the rock. In addition, the predicted *BI* values by ANN, along with their actual values for train and test datasets, are displayed in Figure 9. According to obtained results of this section, R^2 values of 0.75 and 0.85 for train and test stages, respectively showed that the ANN model is able to provide acceptable level of accuracy specially in testing datasets for estimation of the *BI*. ANN is able to determine the importance values of the use inputs in the system (Figure 10). As a result, R_n and V_p are the most important and least important parameters on the *BI* which results of R_n is same as the RF analysis part.

3.2.3. KNN Model

In developing KNN model, several assumptions and parameters were considered. The KNN model was developed to establish the balance between speed and accuracy. Therefore, the model automatically selected the best number of neighbors, within a small range. In the present study, we used k number between the values 3-5 by implementing a trial-and-error method of the system. In addition, the distance computation was based on Euclidean metric. Predicted *BI* values by KNN, along with their actual values for training and testing datasets, are displayed in Figure 11. With R^2 of 0.81 and 0.84 for train and test stages, respectively, in fact, the KNN model is able to offer a balance range for these stages compared to RF and ANN. Figure 12 shows a suggested structure of the KNN predictive model in predicting *BI*.

This figure shows the relationship between the predictors and K selection. In the horizontal axis of the chart, the numbers of the nearest neighbor are displayed. Sums of square errors are shown in the vertical axis. As shown by the figure, the errors for k = 3, 4, and 5 were determined as 372.31, 363.70, and 365.92 respectively. The results revealed that k = 4 is the best value of the nearest neighbor numbers for the developed KNN model. The KNN model also identified the importance of the input

variables (Figure 13). As can be seen, the R_n was identified as the most important variable and followed by V_p, D, and Is_{50}, respectively. It should be noted that R_n was introduced by all RF, ANN and KNN model as the most influential factor on rock *BI*.
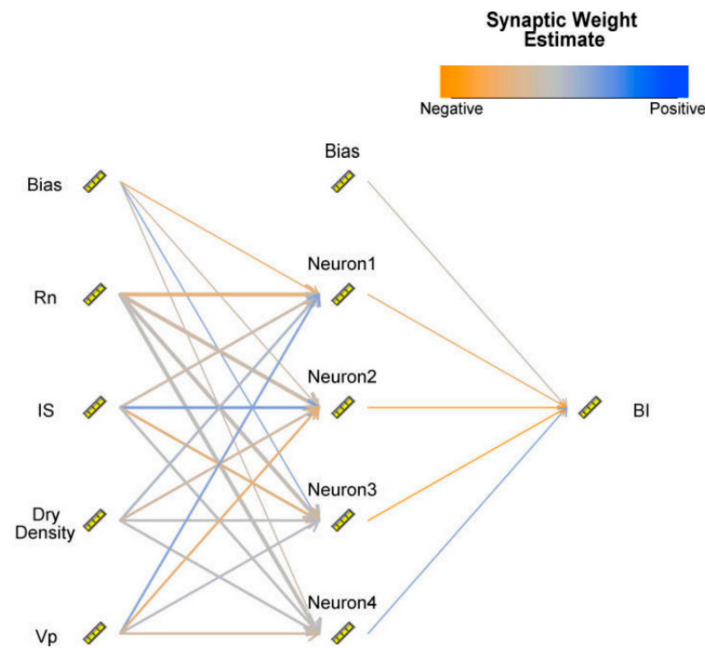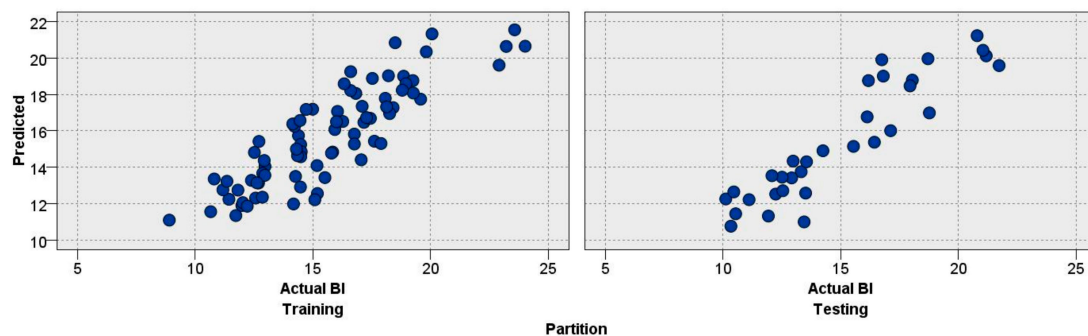

Figure 8. ANN network for predicting the *BI*.

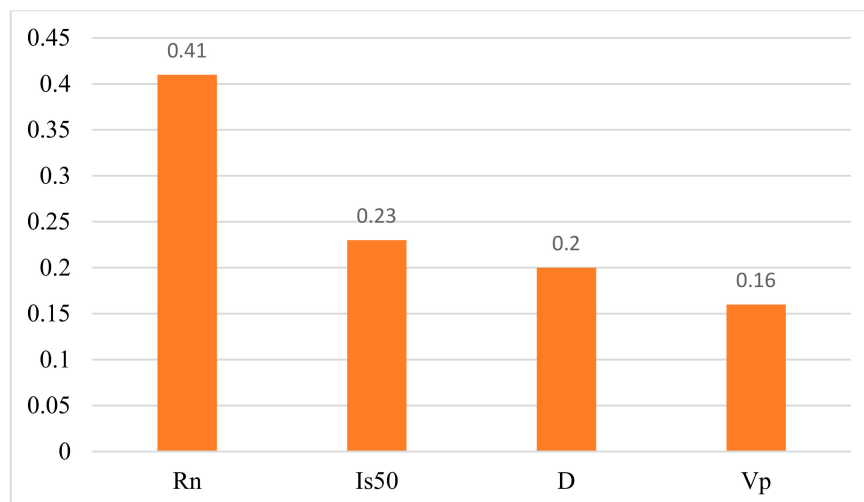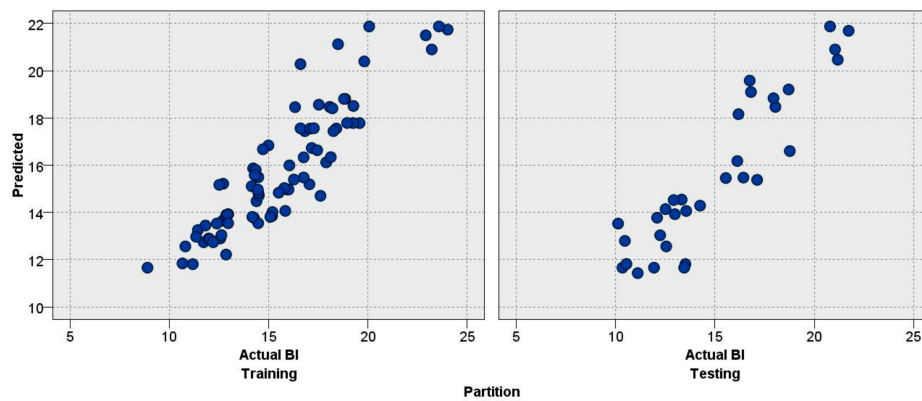**Figure 9.** Testing and training results of ANN model to predict the *BI*.

Figure 10. Input variable's importance to predict the *BI* derived from the ANN model.

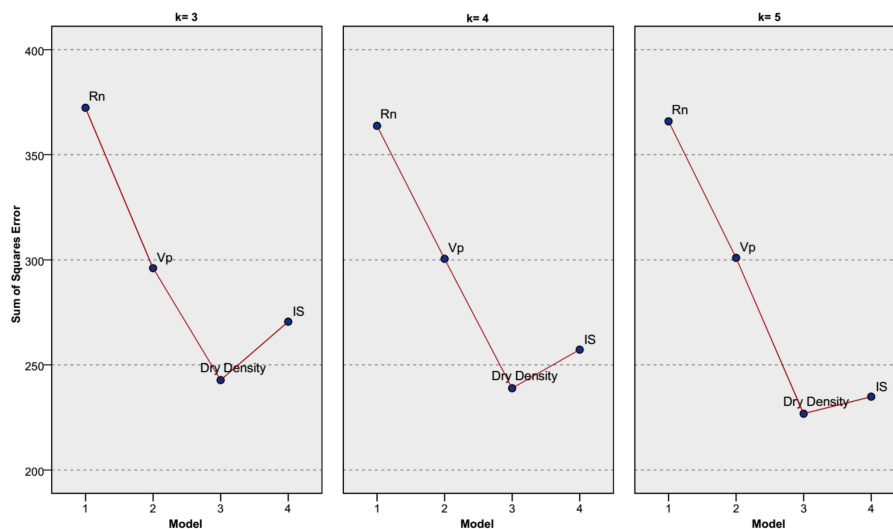**Figure 11.** Testing and training results of KNN model to predict the *BI*.

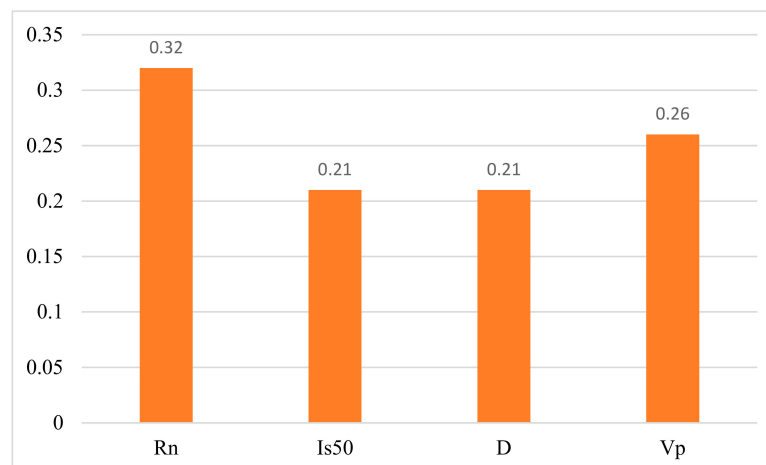**Figure 12.** The relationship between the predictors and K selection for predicting the *BI*.

Figure 13. Input variable's importance to predict the *BI* derived from the KNN model.

4. Validation of the Selected Models

After developing the models for predicting *BI* of the rock, they should be validated through the use of new datasets. Therefore, the authors decided to use 15 more empirical data from the same case study. In should be noted that these data were not used for training and testing phases. The authors used the selected models in previous section and run them using the new datasets for validation

purposes. Then, the measured and predicted values of *BI* were evaluated considering the previous performance indices. Table 3 presents the results of performance indices for all 3 predictive models i.e., ANN, KNN and RF. According to this table, *a20-index* is 1 for all predictive models which shows that m20 (values of experimental/predicted) is equal to N (total number of samples). It confirms that all models are able to provide good results for similar data as well. In addition, R^2 results (0.971, 0.860, and 0.807) and *VAF* results of (96.852, 85.633, and 80.642 %) were obtained for RF, ANN and KNN models, respectively in validation stage which indicate that RF model is better than the other 2 models. In terms of system error, RF model with *RMSE* of 0.62 and *MAE* of 0.46 received lower amount of error compared to ANN and KNN models. Figure 14 shows the measured and predicted *BI* values for the RF, ANN, and KNN models in validation phase. In addition, Figure 15 depicts predicted *BI* values by RF, ANN and KNN together with their measured *BI* for all 15 data samples assigned for validation stage. As it can be seen from these two figures, the *BI* values by RF model are closer to the measured *BI* values in comparison with the KNN and ANN models.

　　　As conclusion on this part, all models are able to provide a good prediction results for *BI* values when similar data will be available. However, RF can receive higher performance capacity if similar data will be available compared to ANN and KNN models. This means that if the other researchers or designers can collect/measure the inputs of this study within their ranges and their properties, it can be expected that the developed RF model is able to predict *BI* values with high correlations and low system error. Therefore, the developed RF model and its structure can be utilized to estimate *BI* of the rock in preliminary design of geotechnical projects subjected to rock mass.
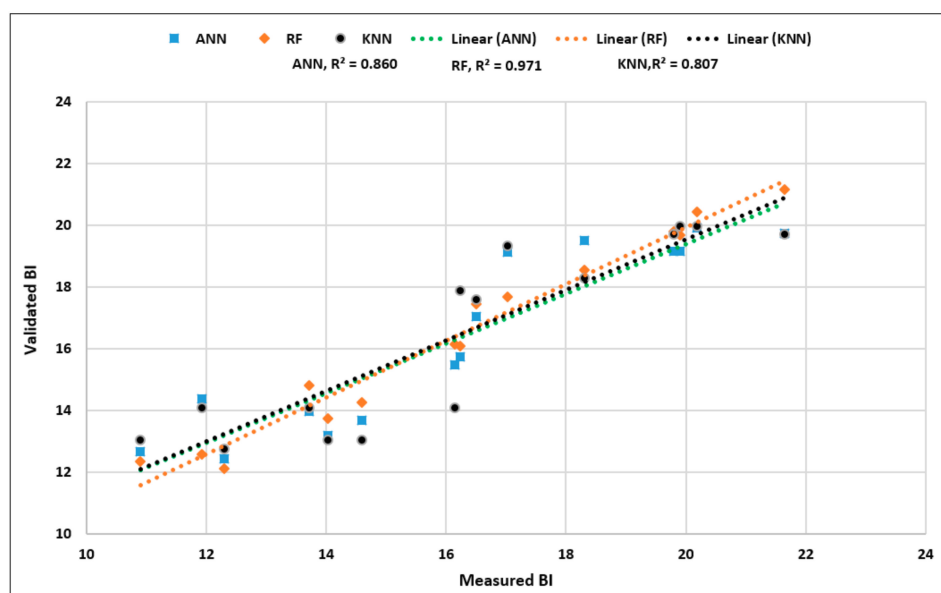

Figure 14. Actual and predicted values for the models selected in validation phase.

Table 3. Performance Assessment for the Validation Phase.

Performance Index	RF	ANN	KNN
R^2	0.971	0.860	0.807
RMSE	0.62	1.22	1.42
VAF (%)	96.852	85.633	80.64
MAE	0.46	0.99	1.14
a20-index	1.00	1.00	1.00

Perfect $R^2 = 1$; Perfect *RMSE* = 0; Perfect *VAF* = 100%; Perfect *MAE* = 0; *a20-index* = 1. Training dataset = TR; Testing dataset = TE.
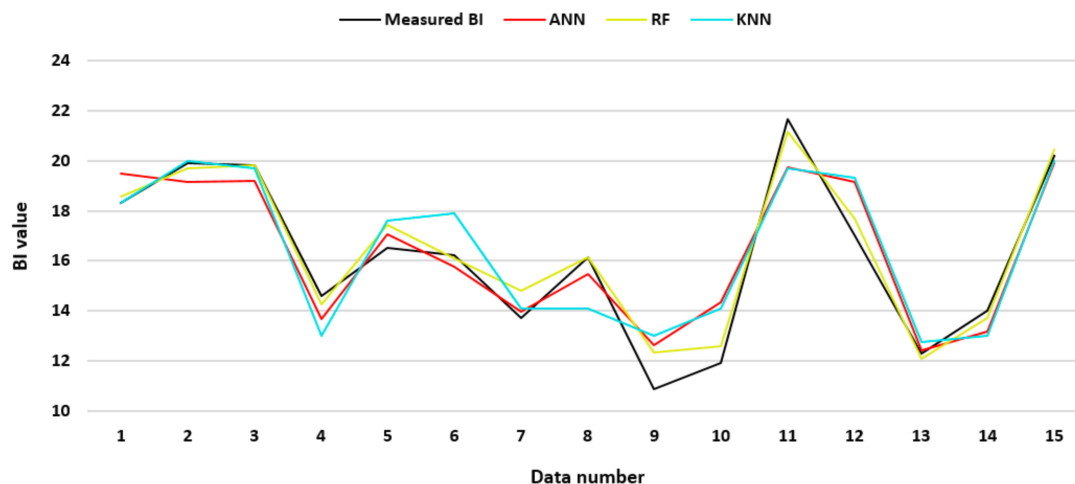
Figure 15. Predicted *BI* values by RF, ANN and KNN together with their measured *BI* for all 15 data samples.

5. Discussion and Conclusions

This present study investigated the application of multiple ML techniques for predicting the rock *BI* using a dataset from a water transfer tunnel in Malaysia. The main aim of this study was to identify the best model (s) in terms of accuracy for both train and test stages. To compare the models, five performance indices, a ranking system, and a gain chart were used. Therefore, five ML models, including the RF, CHAID, ANN, KNN, and SMM were developed. While the results of performance indices showed that the RF outperformed other models for the training dataset, ANN achieved the best ranking for testing dataset. However, the KNN achieved the highest cumulative ranking. A possible explanation for this is that the KNN showed a stable behavior for train and test stages, while the RF and ANN resulted in too many different rankings for train and test stages. Concerning the importance of predictors in this study, all three models, RF, KNN, and ANN identified R_n as the most important factor for predicting the *BI*. The KNN and ANN considered *D* as an important predictor, while the RF did not. This can be explained by the fact that the data of *D* diverged from the average value. It also showed that the RF is intolerable to the dispersion of data points in a data series around the mean.

The RF method outperformed the single-tree based methods like CHAID for both the training and testing stages. The power of RF stems from its abilities to bag a huge number of single tree models and produce an ensemble tree. For categorical data, the RF produces a number of rules which can show the relationships between the predictors and the target variable. However, in this study, the target variable was continuous and the RF could not create a set of rules. While the ANN showed an acceptable performance to predict the *BI*, this method is viewed as black-box. Thus, while it can predict the *BI*, studying its structure does not provide an understanding on the structure of the function being measured. The future studies on the *BI* should cautiously use the KNN. While this method is an intuitive approach and immune to outliers on the predictors [25,63], this may be vulnerable to irrelevant features and correlated inputs [64]. In addition, the ability of KNN to deal with data of mixed-types is still doubtful [25,64].

The last analysis section of this study was related to validation of the selected predictive models i.e., ANN, KNN and RF. To this end, 15 datasets with the same input parameters were considered and then, the ANN, KNN and RF models were run again using these 15 datasets. The results of validation stage showed that RF with R^2 of 0.971 is more capable to predict rock *BI* compared to KNN model with R^2 of 0.807 and ANN model with R^2 of 0.860. This indicates that all models can be used for similar conditions in the future. More specifically, this research suggests to use RF and KNN models (or each of them) by the other researcher or designers in order to predict rock *BI* in design stage of geotechnical projects.

References

1. Miskimins, J.L. The impact of mechanical stratigraphy on hydraulic fracture growth and design considerations for horizontal wells. *Bulletin* **2012**, *91*, 475–499.

2. Rybacki, E.; Reinicke, A.; Meier, T.; Makasi, M.; Dresen, G. What controls the mechanical properties of shale rocks?–Part I: Strength and Young's modulus. *J. Pet. Sci. Eng.* **2015**, *135*, 702–722. [CrossRef]

3. Hajiabdolmajid, V.; Kaiser, P. Brittleness of rock and stability assessment in hard rock tunneling. *Tunn. Undergr. Space Technol.* **2003**, *18*, 35–48. [CrossRef]

4. Kidybiński, A. Bursting liability indices of coal. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* **1981**, *18*, 295–304. [CrossRef]

5. Singh, S.P. Brittleness and the mechanical winning of coal. *Min. Sci. Technol.* **1986**, *3*, 173–180. [CrossRef]

6. Singh, S.P. Burst energy release index. *Rock Mech. Rock Eng.* **1988**, *21*, 149–155. [CrossRef]

7. Yagiz, S. Utilizing rock mass properties for predicting TBM performance in hard rock condition. *Tunn. Undergr. Space Technol.* **2008**, *23*, 326–339. [CrossRef]

8. Yagiz, S.; Gokceoglu, C. Application of fuzzy inference system and nonlinear regression models for predicting rock brittleness. *Expert Syst. Appl.* **2010**, *37*, 2265–2272. [CrossRef]

9. Altindag, R. Reply to the Discussion by Yagiz on "Assessment of Some Brittleness Indexes in Rock-Drilling Efficiency" by Altindag. *Rock Mech. Rock Eng.* **2010**, *43*, 375–376. [CrossRef]

10. Yagiz, S. Assessment of brittleness using rock strength and density with punch penetration test. *Tunn. Undergr. Space Technol.* **2009**, *24*, 66–74. [CrossRef]

11. Morley, A. *Strength of Material*; Longmans: Suffolk, UK, 1944.

12. Ramsay, J.G. Folding and fracturing of rocks. *McGraw Hill B Co.* **1967**, *568*, 289.

13. Obert, L.; Duvall, W.I. *Rock Mechanics and the Design of Structures in Rock*; John Wiley & Sons Inc.: Hoboken, NJ, USA, 1967; Volume 278.

14. Altindag, R.; Guney, A. Predicting the relationships between brittleness and mechanical properties (UCS, TS and SH) of rocks. *Sci. Res. Essays* **2010**, *5*, 2107–2118.

15. Wang, Y.; Watson, R.; Rostami, J.; Wang, J.Y.; Limbruner, M.; He, Z. Study of borehole stability of Marcellus shale wells in longwall mining areas. *J. Pet. Explor. Prod. Technol.* **2014**, *4*, 59–71. [CrossRef]

16. Koopialipoor, M.; Noorbakhsh, A.; Noroozi Ghaleini, E.; Jahed Armaghani, D.; Yagiz, S. A new approach for estimation of rock brittleness based on non-destructive tests. *Nondestruct. Test. Eval.* **2019**, 1–22. [CrossRef]

17. Khandelwal, M.; Faradonbeh, R.S.; Monjezi, M.; Armaghani, D.J.; Majid, M.Z.B.A.; Yagiz, S. Function development for appraising brittleness of intact rocks using genetic programming and non-linear multiple regression models. *Eng. Comput.* **2017**, *33*, 13–21. [CrossRef]

18. Nejati, H.R.; Moosavi, S.A. A new brittleness index for estimation of rock fracture toughness. *J. Min. Environ.* **2017**, *8*, 83–91.

19. Hajihassani, M.; Abdullah, S.S.; Asteris, P.G.; Armaghani, D.J. A Gene Expression Programming Model for Predicting Tunnel Convergence. *Appl. Sci.* **2019**, *9*, 4650. [CrossRef]

20. Asteris, P.G.; Armaghani, D.J.; Hatzigeorgiou, G.D.; Karayannis, C.G.; Pilakoutas, K. Predicting the shear strength of reinforced concrete beams using Artificial Neural Networks. *Comput. Concr.* **2019**, *24*, 469–488.

21. Zhou, J.; Bejarbaneh, B.Y.; Armaghani, D.J.; Tahir, M.M. Forecasting of TBM advance rate in hard rock condition based on artificial neural network and genetic programming techniques. *Bull. Eng. Geol. Environ.* **2019**. [CrossRef]

22. Yong, W.; Zhou, J.; Armaghani, D.J.; Tahir, M.M.; Tarinejad, R.; Pham, B.T.; Van Huynh, V. A new hybrid simulated annealing-based genetic programming technique to predict the ultimate bearing capacity of piles. *Eng. Comput.* **2020**. [CrossRef]

23. Zhou, J.; Guo, H.; Koopialipoor, M.; Armaghani, D.J.; Tahir, M.M. Investigating the effective parameters on the risk levels of rockburst phenomena by developing a hybrid heuristic algorithm. *Eng. Comput.* **2020**. [CrossRef]

24. Mahdiyar, A.; Jahed Armaghani, D.; Koopialipoor, M.; Hedayat, A.; Abdullah, A.; Yahya, K. Practical Risk Assessment of Ground Vibrations Resulting from Blasting, Using Gene Expression Programming and Monte Carlo Simulation Techniques. *Appl. Sci.* **2020**, *10*, 472. [CrossRef]

25. Zhou, J.; Li, X.; Mitri, H.S. Evaluation method of rockburst: State-of-the-art literature review. *Tunn. Undergr. Space Technol.* **2018**, *81*, 632–659. [CrossRef]

26. Zhou, J.; Li, E.; Yang, S.; Wang, M.; Shi, X.; Yao, S.; Mitri, H.S. Slope stability prediction for circular mode failure using gradient boosting machine approach based on an updated database of case histories. *Saf. Sci.* **2019**, *118*, 505–518. [CrossRef]

27. Zhou, J.; Shi, X.; Li, X. Utilizing gradient boosted machine for the prediction of damage to residential structures owing to blasting vibrations of open pit mining. *J. Vib. Control* **2016**, *22*, 3986–3997. [CrossRef]

28. Zhou, J.; Li, X.; Mitri, H.S. Comparative performance of six supervised learning methods for the development of models of hard rock pillar stability prediction. *Nat. Hazards* **2015**, *79*, 291–316. [CrossRef]

29. Yang, H.; Liu, J.; Liu, B. Investigation on the cracking character of jointed rock mass beneath TBM disc cutter. *Rock Mech. Rock Eng.* **2018**, *51*, 1263–1277. [CrossRef]

30. Yang, H.Q.; Li, Z.; Jie, T.Q.; Zhang, Z.Q. Effects of joints on the cutting behavior of disc cutter running on the jointed rock mass. *Tunn. Undergr. Space Technol.* **2018**, *81*, 112–120. [CrossRef]

31. Chen, H.; Asteris, P.G.; Jahed Armaghani, D.; Gordan, B.; Pham, B.T. Assessing Dynamic Conditions of the Retaining Wall: Developing Two Hybrid Intelligent Models. *Appl. Sci.* **2019**, *9*, 1042. [CrossRef]

32. Liu, B.; Yang, H.; Karekal, S. Effect of Water Content on Argillization of Mudstone during the Tunnelling process. *Rock Mech. Rock Eng.* **2019**. [CrossRef]

33. Armaghani, D.J.; Hatzigeorgiou, G.D.; Karamani, C.; Skentou, A.; Zoumpoulaki, I.; Asteris, P.G. Soft computing-based techniques for concrete beams shear strength. *Procedia Struct. Integr.* **2019**, *17*, 924–933. [CrossRef]

34. Apostolopoulou, M.; Armaghani, D.J.; Bakolas, A.; Douvika, M.G.; Moropoulou, A.; Asteris, P.G. Compressive strength of natural hydraulic lime mortars using soft computing techniques. *Procedia Struct. Integr.* **2019**, *17*, 914–923. [CrossRef]

35. Xu, H.; Zhou, J.G.; Asteris, P.; Jahed Armaghani, D.; Tahir, M.M. Supervised Machine Learning Techniques to the Prediction of Tunnel Boring Machine Penetration Rate. *Appl. Sci.* **2019**, *9*, 3715. [CrossRef]

36. Huang, L.; Asteris, P.G.; Koopialipoor, M.; Armaghani, D.J.; Tahir, M.M. Invasive Weed Optimization Technique-Based ANN to the Prediction of Rock Tensile Strength. *Appl. Sci.* **2019**, *9*, 5372. [CrossRef]

37. Asteris, P.G.; Mokos, V.G. Concrete compressive strength using artificial neural networks. *Neural Comput. Appl.* **2019**. [CrossRef]

38. Asteris, P.G.; Nikoo, M. Artificial bee colony-based neural network for the prediction of the fundamental period of infilled frame structures. *Neural Comput. Appl.* **2019**. [CrossRef]

39. Asteris, P.G.; Moropoulou, A.; Skentou, A.D.; Apostolopoulou, M.; Mohebkhah, A.; Cavaleri, L.; Rodrigues, H.; Varum, H. Stochastic Vulnerability Assessment of Masonry Structures: Concepts, Modeling and Restoration Aspects. *Appl. Sci.* **2019**, *9*, 243. [CrossRef]

40. Kaunda, R.B.; Asbury, B. Prediction of rock brittleness using nondestructive methods for hard rock tunneling. *J. Rock Mech. Geotech. Eng.* **2016**, *8*, 533–540. [CrossRef]

41. Kass, G.V. An exploratory technique for investigating large quantities of categorical data. *J. R. Stat. Soc. Ser. C Appl. Stat.* **1980**, *29*, 119–127. [CrossRef]

42. Brown, G. Ensemble Learning. In *Encyclopedia of Machine Learning*; Springer: Boston, MA, USA, 2010; pp. 312–320.

43. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

44. Wu, X.; Kumar, V.; Quinlan, J.R.; Ghosh, J.; Yang, Q.; Motoda, H.; McLachlan, G.J.; Ng, A.; Liu, B.; Philip, S.Y. Top 10 algorithms in data mining. *Knowl. Inf. Syst.* **2008**, *14*, 1–37. [CrossRef]

45. Akbulut, Y.; Sengur, A.; Guo, Y.; Smarandache, F. NS-k-NN: Neutrosophic set-based k-nearest neighbors classifier. *Symmetry* **2017**, *9*, 179. [CrossRef]

46. Qian, Y.; Zhou, W.; Yan, J.; Li, W.; Han, L. Comparing machine learning classifiers for object-based land cover classification using very high resolution imagery. *Remote Sens.* **2015**, *7*, 153–168. [CrossRef]

47. Kavzoglu, T.; Sahin, E.K.; Colkesen, I. Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. *Landslides* **2014**, *11*, 425–439. [CrossRef]

48. Hasanipanah, M.; Monjezi, M.; Shahnazar, A.; Armaghani, D.J.; Farazmand, A. Feasibility of indirect determination of blast induced ground vibration based on support vector machine. *Measurement* **2015**, *75*, 289–297. [CrossRef]

49. Kalantar, B.; Pradhan, B.; Naghibi, S.A.; Motevalli, A.; Mansor, S. Assessment of the effects of training data selection on the landslide susceptibility mapping: A comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN). *Geomat. Nat. Hazards Risk* **2018**, *9*, 49–69. [CrossRef]

50. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]

51. Hong, H.; Pradhan, B.; Bui, D.T.; Xu, C.; Youssef, A.M.; Chen, W. Comparison of four kernel functions used in support vector machines for landslide susceptibility mapping: A case study at Suichuan area (China). *Geomat. Nat. Hazards Risk* **2017**, *8*, 544–569. [CrossRef]

52. Kamavisdar, P.; Saluja, S.; Agrawal, S. A survey on image classification approaches and techniques. *Int. J. Adv. Res. Comput. Commun. Eng.* **2013**, *2*, 1005–1009.

53. Liu, J.; Savenije, H.H.G.; Xu, J. Forecast of water demand in Weinan City in China using WDF-ANN model. *Phys. Chem. Earth Parts A/B/C* **2003**, *28*, 219–224. [CrossRef]

54. Mohamad, E.T.; Armaghani, D.J.; Noorani, S.A.; Saad, R.; Alvi, S.V.; Abad, N.K. Prediction of flyrock in boulder blasting using artificial neural network. *Electron. J. Geotech. Eng.* **2012**, *17*, 2585–2595.

55. Tonnizam Mohamad, E.; Hajihassani, M.; Jahed Armaghani, D.; Marto, A. Simulation of blasting-induced air overpressure by means of Artificial Neural Networks. *Int. Rev. Model. Simul.* **2012**, *5*, 2501–2506.

56. Asteris, P.; Roussis, P.; Douvika, M. Feed-forward neural network prediction of the mechanical properties of sandcrete materials. *Sensors* **2017**, *17*, 1344. [CrossRef] [PubMed]

57. Apostolopoulour, M.; Douvika, M.G.; Kanellopoulos, I.N.; Moropoulou, A.; Asteris, P.G. Prediction of Compressive Strength of Mortars using Artificial Neural Networks. In Proceedings of the 1st International Conference TMM_CH, Transdisciplinary Multispectral Modelling and Cooperation for the Preservation of Cultural Heritage, Athens, Greece, 10–13 October 2018; pp. 10–13.

58. Asteris, P.G.; Argyropoulos, I.; Cavaleri, L.; Rodrigues, H.; Varum, H.; Thomas, J.; Lourenço, P.B. Masonry compressive strength prediction using artificial neural networks. In *International Conference on Transdisciplinary Multispectral Modeling and Cooperation for the Preservation of Cultural Heritage*; Springer: Cham, Germany, 2018; pp. 200–224.

59. Armaghani, D.J.; Mohamad, E.T.; Narayanasamy, M.S.; Narita, N.; Yagiz, S. Development of hybrid intelligent models for predicting TBM penetration rate in hard rock condition. *Tunn. Undergr. Space Technol.* **2017**, *63*, 29–43. [CrossRef]

60. Momeni, E.; Nazir, R.; Armaghani, D.J.; Maizir, H. Application of artificial neural network for predicting shaft and tip resistances of concrete piles. *Earth Sci. Res. J.* **2015**, *19*, 85–93. [CrossRef]

61. Ulusay, R.; Hudson, J.A. *ISRM (2007) The Complete ISRM Suggested Methods for Rock Characterization, Testing and Monitoring: 1974–2006*; International Society for Rock Mechanics, Commission on Testing Methods: Ankara, Turkey, 2007; p. 628.

62. Hucka, V.; Das, B. Brittleness determination of rocks by different methods. *Int. J. Rock Mech. Min. Sci. Geomech. Abstr.* **1974**, *11*, 389–392. [CrossRef]

63. Salaria, S.; Drozd, A.; Podobas, A.; Matsuoka, S. Predicting performance using collaborative filtering. In Proceedings of the 2018 IEEE International Conference on Cluster Computing (CLUSTER), Belfast, UK, 10–13 September 2018; pp. 504–514.

64. Su, G.S.; Zhang, X.F.; Yan, L.B. Rockburst prediction method based on case reasoning pattern recognition. *J. Min. Saf. Eng.* **2008**, *1*, 15.