DOI : https://doi.org/10.7186/wg451201903

Predicting uniaxial compressive strength using Support Vector Machine algorithm

Hafedz Zakaria[1], Rini Asnida Abdullah[2,*], Amelia Ritahani Ismail[3], Mohd For Mohd Amin[2]

[1]Public Works Department of Malaysia, Jalan Sultan Salahuddin, 50582 Kuala Lumpur, Malaysia
[2]Faculty of Civil Engineering, Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia
[3]Kulliyah of ICT, International Islamic University, 50728 Kuala Lumpur, Malaysia
[*]Corresponding author email address: asnida@utm.my

Abstract: Compressive strength is the most important parameter in rock since all loads will be transferred and rest on the rock which is based on the load bearing capacity of rock in compression. However, obtaining the compressive strength or mostly measured, the uniaxial compressive strength (UCS) from the laboratory test requires certain standard and also cost constrain. This paper presents the application of Support Vector Machine (SVM) algorithm to predict the UCS. An algorithm has been tested on a series of rock data using dry density and velocity parameters. The relationship between the dry density, sonic velocity, and UCS was analyzed using RapidMiner Studio software. From the result, it was found that SVM is capable of predicting the missing values with a prediction trend accuracy of 75%. The results obtained and observation made in this study suggests that SVM could be a reliable tool to predict the UCS of a given rock. More robust prediction can be established with bigger sample number. It is worth mentioning, that the program module that has been set up could be used repeatedly for other correlation problems.

Keywords: unconfined compressive strength, dry density, sonic wave velocity, support vector machine

INTRODUCTION

Rock mechanics is an essential part of geotechnical engineering. One of the most important parameters of rocks is the uniaxial compressive strength of the rock sample. However, obtaining sufficient core sample with the proper quality for direct testing is often hindered especially in weathered or/and fractured rocks. Most commonly used correlation is using the sonic wave velocity logging, which is an indirect test. The correlation established by McNally, 1990 states that, UCS:

$$UCS = 1000 \times e^{-0.035t} \quad \text{(Equation 1)}$$

where the UCS is in MPa and t is the sonic travel time in microsecond/feet.

However, some researchers have suggested that this equation does not take into account the variations due to rock mass parameters (Hatherly *et al.*, 2009; Medhurst *et al.*, 2010; Barton, 2006). Other researchers suggested that the correlation of UCS and sonic velocity should be done on a regional basis, and they also believe that a linear relationship is the most appropriate for this relationship (Sharma & Singh, 2008).

In establishing the correlation, firstly, the regression analysis techniques are often used to predict unavailable data. This technique has a major flaw, of which it only predicts the mean value of the data. This may lead to inaccurate prediction, especially in non-linear equation problems. This flaw may be countered using supervised machine learning algorithm; machine learning is part of computer science, developed from the study of pattern recognition and computational learning theory in Artificial Intelligence. Machine learning explores the study and construction of algorithms that can learn from and make predictions on data. Commonly known machine learning is Artificial Neural Networks (ANN). Regardless, the ANN does not force the predicted data to be a mean value, thus preserving and using the existing variance of the measured data.

This limitation has been resolved using the SVM. Therefore, this study aims to predict missing UCS values using sonic wave velocity and dry density parameters by utilizing the SVM.

LITERATURE REVIEW

Support vector algorithm is a nonlinear generalization of the *Generalized Portrait* algorithm developed in Russia in the sixties (Vapnik & Lerner, 1963; Vapnik & Chervonenkis, 1964). It is a supervised learning model with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis. In regression and time series prediction applications, excellent performances were reported using SVM (Muller *et al.*, 1997; Drucker *et al.*, 1997; Stitson *et al.*, 1999; Mattera & Haykin, 1999). Just recently, SVM has come into attention for solving the engineering problem, in which SVM characterizes properties of learning machines which enable them to generalize well to unseen data.

SVMs can be used to solve in many real life problem such as 1) classification of images, 2) hand-written characters recognition and 3) predictions.

EXPERIMENTAL SETUP

In establishing the results, the experiment is divided into data preparation phase, which is described in later section.

Data preparation

In this study, a total of 160 data, of which 140 data is regarded as complete (ie: have both UCS and sonic wave velocity values) of which are used to learn, and 20 data that are incomplete has been gathered from the Jabatan Kerja Raya (JKR). The data is then analyzed and cleansed to get rid of unwanted attributes such as the site location, borehole number, sample number, etc. The attribute chosen to be analyzed is trimmed and copied into another file. Two files were created through this process. One of the file consists of complete data set (UCS, sonic velocity and dry density), whilst the other file consists of data set with missing UCS values. The first file (learning data set) is important to train the SVM algorithm the existing correlation values between the attributes so that it can be applied and predicted the missing values in the second file (prediction data set). The data is then imported into the Rapid Miner Studio software, of which shall be explained in the following sections.

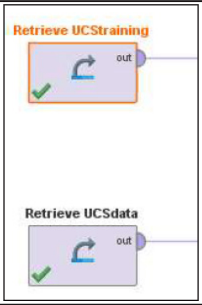
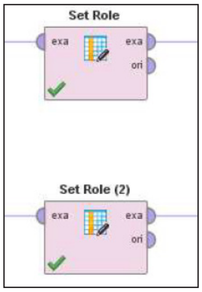
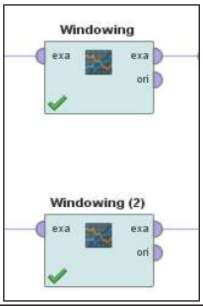
Programming the module

Rapid Miner software comes with a user-friendly graphical user interface (GUI). However, to use the software effectively, the function and limitation of each operator must be understood before it can be utilized. Each of the operators will be explained in details in the Table 1.

RESULTS AND DISCUSSION

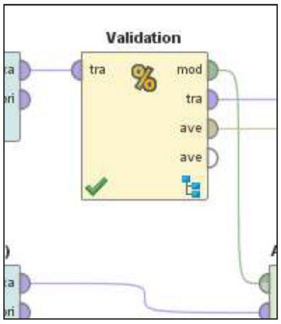
Once the analysis is completed, the RapidMiner software; thru the Forecasting Performance operator, gave a prediction trend accuracy of 75%, which is accurate enough considering that there are not enough data as compared to a major research involving thousands of data. Figure 1 shows almost identical trend between the actual UCS versus the predicted UCS values shown almost identical trend, in which reflects to the good agreement found between these values.

Table 1: SVM operating module.

Operator	Screenshot
Retrieve operator This operator reads an object from the data repository. It is more efficient to use this operator since it provides full meta data processing, which eases the usage of RapidMiner a lot. In contrast to accessing a raw file, it provides the complete meta data of the data set, so all meta data transformations are possible.	
Set Role operator This operator is used to change the role of one or more attributes. Changing the role of an attribute may change the part played by that attribute in a process. Different learning operators require attributes with different roles. This operator is used to set the right roles for attributes before applying the desired operator. At this stage, the UCS attribute are set to the "label" role as it acts as a target attribute for learning operators (e.g. SVM, Decision Tree, etc.).At the bottom Set Role operator, the UCS attribute is set to a prediction role, as it will act as a predicted attribute of a learning scheme.	
Windowing operator This operator transforms a given example set containing series data into a new example set containing single valued examples. For this purpose, windows with a specified window and step size are moved across the series and the attribute value lying horizon values after the window end is used as label which should be predicted.	

<table>
<tr>
<td>
X-Validation operator

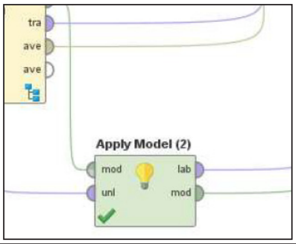
This operator performs a cross-validation in order to estimate the statistical performance of a learning operator. It is mainly used to estimate how accurately a model will perform in practice.

There are two sub-processes in this operator: a training subprocess and a testing subprocess. The training subprocess is used for training a model. The trained model is then applied in the testing subprocess. The performance of the model is also measured during the testing phase.
</td>
<td>


</td>
</tr>
<tr>
<td>
SVM operator

Nested inside the X-Validation operator, at the training subprocess, is the SVM (Support Vector Machine) Learner operator. This learning method can be used for both regression and classification and provides a fast algorithm and good results for many learning tasks. This operator supports various kernel types including dot, radial, polynomial, neural, anova, epachnenikov, Gaussian combination and multiquadric.

The standard SVM takes a set of input data and predicts, for each given input, which of the two possible classes comprises the input, making the SVM a non-probabilistic binary linear classifier. Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other. An SVM model is representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on.
</td>
<td></td>
</tr>
<tr>
<td>
Apply Model operator

This operator applies an already learnt or trained model on a dataset. After the model is trained using learning operators (eg: SVM, Decision Tree, etc.), the model can be applied on another data set; usually for prediction purposes. All needed parameters are stored within the model object. It is compulsory that both data sets have exactly the same number, order, type and role of attributes. If these properties of meta data are not consistent, it may lead to serious errors.
</td>
<td>


</td>
</tr>
<tr>
<td>
Forecasting Performance operator

This operator delivers as output a list of performance values according to a list of selected performance criteria for forecasting regression tasks.
</td>
<td></td>
</tr>
</table>

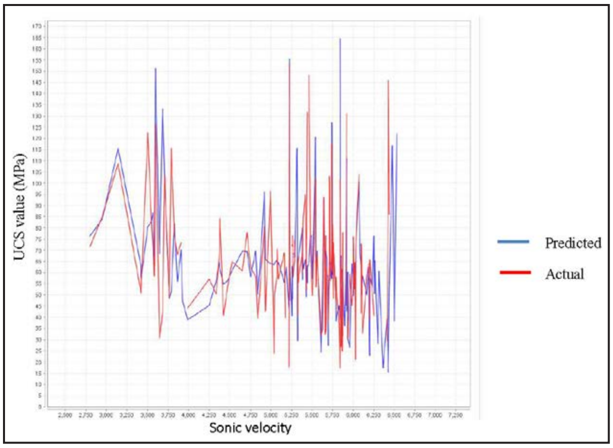


Figure 1: UCS (predicted and actual) versus sonic velocity.

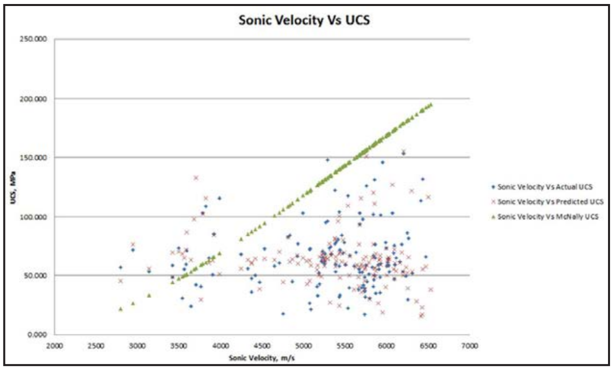


Figure 2: UCS versus sonic velocity.

A comparison has been made with the actual, predicted and UCS value that have been obtained using McNally's equation (Equation 1). Figure 2 indicates that the predicted UCS values are closer to the actual UCS values compared to the UCS values generated by the McNally's equation. This confirmed that the predicted data will yield a more accurate result as compared to the existing equation.

It should be highlighted that, this situation can be very helpful in cases of projects that rely on small existing data to predict any missing values that are encountered.

CONCLUSIONS

The results obtained and the observation made in this study draw some conclusions. Support Vector Machine algorithm has been successfully utilized to predict the UCS value with 75% accuracy. A more robust prediction can be achieved with bigger number of data set. It is worth mentioning that, the programmed module that was generated can be used to predict any missing data that has an established correlation. This could make future prediction easier and faster. Most importantly, there are still many application of SVM that can and must be explored to solve the engineering difficulties.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the Ministry of Higher Education Malaysia (MOHE) under Q.J130000.2522.19H81 grant.

REFERENCES

Barton, N., 2006. Relationships between rock quality, depth and seismic velocity. In: Rock Quality, Seismic Velocity, Attenuation and Anisotropy, Taylor and Francis Group, London, p. 69-96.

Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A. & Vapnik, V., 1997. Support vector regression machines. Advances in Neural Information Processing Systems, 9, Cambridge, MA. MIT Press, 155-161.

Hatherly, P., Medhurst, T., Ye, G. & Payne, D., 2009. Geotechnical evaluation of roof conditions at Crinum Mine based on geophysical log interpretation. In: N. Aziz (Ed.), Proceedings Coal 2009: Coal Operators' Conference, Wollongong Australia, Australasian Institute of Mining and Metallurgy. http://ro.uow.edu.au/coal/67/.

Mattera, D. & Haykin, S., 1999, Support vector machines for dynamic reconstruction of a chaotic system. Advances in Kernel Methods – Support Vector Learning, Cambridge, MA. MIT Press, 211-242.

McNally, G.H., 1990. The Prediction of Geotechnical Rock Properties from Sonic and Neutron Logs. Exploration Geophysics, 21, 65-71.

Medhurst, T., Hatherly, P. & Zhou, B., 2010. 3D geotechnical models for coal and clastic rocks based on the GSR. In: Proceedings of the 10th Coal Operators' Conference, Wollongong, 16-20 February. http://ro.uow.edu.au/coal/301/.

Muller, K.R., Smola, A., Ratsch, G., Scholkopf, B., Kohlmorgen, J. & Vapnik, V., 1997. Predicting time series with support vector machines. Artificial Neural Networks – ICANN'97, Berlin, 999-1004.

Sharma, P. & Singh, T., 2008. A correlation between P-wave velocity, impact strength index, slake durability index and uniaxial compressive strength. Bulletin of Engineering Geology and the Environment, 67(1), 17.

Stitson, M., Gammerman, A., Vapnik, V., Vovk, V., Watkins, C. & Weston, J., 1999. Support vector regression with ANOVA decomposition kernels. Advances in Kernel Methods – Support Vector Learning, Cambridge, MA. MIT Press, 285-292.

Vapnik,V. & Chervonenkis, A., 1964. A note on one class of perceptrons. Automation and Remote Control, 25.

Vapnik, V. & Lerner, A., 1963. Pattern recognition using generalized portrait method. Automation and Remote Control, 24.

Manuscript received 13 October 2017
Revised manuscript received 7 January 2019
Manuscript accepted 28 February 2019