PREDICTION OF TBM PENETRATION RATE USING SUPPORT VECTOR MACHINE

PREVISÃO DA TAXA DE PENETRAÇÃO TBM UTILIZANDO MÁQUINA DE VETOR DE SUporte

PREDICCIÓN DE LA TASA DE PENETRACIÓN DE TBM UTILIZANDO LA MÁQUINA DE VECTORES DE SOPORTE

https://doi.org/10.26895/geosaberes.v11i0.1048

ALIREZA AFRADI 1
ARASH EBRAHIMABADI 2*
TAHEREH HALLAJIAN 3

1 PhD Candidate, Department of Mining and Geology, Qaemshahr Branch, Islamic Azad University, Qaemshahr, Iran. CP: 4765161964, Tel.:(+98) 9385767009, alirezaafradi@yahoo.com, http://orcid.org/0000-0002-1071-3990

2 Associate Professor, Department of Mining and Geology, Qaemshahr Branch, Islamic Azad University, Qaemshahr, Iran. CP: 4765161964, Tel.:(+98) 9123221732, a.ebrahimbadi@qaemiau.ac.ir, http://orcid.org/0000-0002-1996-2731
*Corresponding author

3 Assistant Professor, Department of Mining and Geology, Qaemshahr Branch, Islamic Azad University, Qaemshahr, Iran. CP: 4765161964, Tel.:(+98) 9113912040, hallajian.san@gmail.com, http://orcid.org/0000-0002-2513-9944

ABSTRACT

One of the most important issues in mechanized excavating is to predict the TBM penetration rate. Understanding the factors influencing the rate of penetration is important, which allows for a more accurate estimation of the stopping and excavating times and operating costs. In this study, Input and output parameters including Uniaxial Compressive Strength (UCS), Brazilian Tensile Strength (BTS), Peak Slope Index (PSI), Distance between Planes of Weakness (DPW), Alpha angle and Rate of Penetration (ROP) (m/hr) in the Queens Water Tunnel using support vector machine. Results showed that prediction of penetration rate for Support Vector Machine (SVM) method is \( R^2 = 0.9678 \) and RMSE = 0.064778. According to the results, Support Vector Machine (SVM) is effective and has high accuracy.

Keywords: TBM. Penetration rate. Support Vector Machine (SVM).

RESUMO

Uma das questões mais importantes na escavação mecanizada é prever a taxa de penetração do TBM. É importante compreender os fatores que influenciam a taxa de penetração, o que permite uma estimativa mais precisa dos tempos de parada e escavação e dos custos operacionais. Neste estudo, os parâmetros de entrada e saída incluem resistência à compressão uniaxial (UCS), resistência à tração brasileira (BTS), índice de inclinação de pico (PSI), distância entre planos de fraqueza (DPW), ângulo alfa e taxa de penetração (ROP) (m / hr) no túnel de água de Queens usando máquina de vetor de suporte. Os resultados mostraram que a previsão da taxa de penetração para o método SVM (Support Vector Machine) é \( R^2 = 0,9678 \) e RMSE = 0,064778. De acordo com os resultados, o SVM (Support Vector Machine) é eficaz e tem alta precisão.

Palavras-chave: TBM. Taxa de penetração. Máquina de vetores de suporte (SVM).

RESUMEN

Una de las cuestiones más importantes en la excavación mecanizada es predecir la tasa de penetración de TBM. Es importante comprender los factores que influyen en la velocidad de penetración, lo que permite una estimación más precisa de los tiempos de parada y excavación y los costos operativos. En este estudio, los parámetros de entrada y salida incluyen la...
resistencia a la compresión uniaxial (UCS), la resistencia a la tracción brasileña (BTS), el índice de pendiente máxima (PSI), la distancia entre los planos de debilidad (DPW), el ángulo alfa y la tasa de penetración (ROP) (m / h) en el túnel de agua de Queens utilizando la máquina de vectores de soporte. Los resultados mostraron que la predicción de la tasa de penetración para el método de la máquina de vectores de soporte (SVM) es $R^2 = 0.9678$ y RMSE = 0.064778, según los resultados, la máquina de vectores de soporte (SVM) es efectiva y tiene una alta precisión.

Palabras Clave: TBM. Tasa de penetración. Máquina de vectores de soporte (SVM).

INTRODUCTION

Nowadays, in many major cities around the world, urban transport tunnels play an important role in human life, necessitating the use of modern tools such as tunnel boring machine to excavate and execute these projects (TARKOY, 1973; FARMER, GLOSSOP, 1980; HASSANPOUR et al., 2009). The speed and quality of excavating machines make them competitive with traditional methods (CASSINELLI et al, 1982; YAGIZ, KARAHAN, 2015). One of the important factors is TBM penetration rate (LISLERUD et al., 1983; ARMAGHANI et al., 2019). TBM penetration rate estimates can be used to reduce the risks associated with the costs of common investment in excavating operations (BIENIAWSKI et al., 2007; ZHAO et al., 2019). However, TBMs are sensitive to geological conditions such as fractures and cracks, tunneling is a high-risk industry that increases the risk with the use of mechanized tunnels (HASSANPOUR et al., 2009; HASSANPOUR et al., 2011; YAGIZ, KARAHAN, 2015). Estimating TBM penetration rate has a significant impact on controlling the time and cost of the project and choosing the excavating method (GONG, ZHAO, 2009; ADOKO et al., 2017; VERGARAA, SAROGLOU, 2017). The complexity of the reaction between rock mass and TBM makes it extremely difficult to estimate TBM penetration rate (KHADEMI HAMIDI et al., 2010; ZARE NAGHADEHI et al., 2018). The penetration rate is defined as the ratio of excavating distance to excavating time during a continuous excavating phase (YAGIZ, 2008; YAGIZ et al., 2018).

Penetration rate models used in engineering can be divided into three categories: 1. Experimental Models 2. Theory Models 3. Numerical models

Experimental models are often obtained by analyzing data from tunnel projects (BAMFORD, 1984; BARTON, 2000), while theoretical models are obtained by performing laboratory tests and simulating reality in laboratories (INNAURATO, 1991; RIBACCHI, LEMBO-FAZIO, 2005; ARMETTI et al., 2018), Numerical models that have received much attention in recent years are a new and less expensive method that reflects the reality of using project records (AFRADI et al., 2019).

METHODOLOGY

Support Vector Machine

Support Vector Machine (SVM) is a good and efficient classifier of the Kernel Methods branch of machine learning, which has been particularly well-known in recent years for classification, regression and pattern recognition issues (VAPNIK, 1995). Support Vector Machine (SVM) supervised learning methods. The benefits of Support Vector Machine (SVM) are relatively simple training and, unlike neural networks, do not get stuck in local maxims (VAPNIK, 1998). It also works well for large data sets. Support Vector Machine (SVM) has been used in various sciences such as engineering, medicine, agriculture, natural resources, accounting, management, civil, economics, natural resources, speech recognition, pattern recognition and so on (XU et al., 2019). Some of its applications include pilotless aircraft control, aircraft deflection tracking, welding quality analysis, computer quality analysis, chemical product design analysis, synthetic member design, organ transplant time optimization,
oil and gas exploration, voice recognition, instantaneous language translator, system Client Cash Processing Systems, Truck Brake Detection Systems, Drug Detection, Signature Review, Loan Risk Estimation, Capital Assessment, Market Forecasting, Energy Essentials Forecasting, Pharmaceutical Response Forecasting, Weather Forecasting, Document Inspection, Target Detection, Face Detection, Noise Prevention, Image Recognition, Signal Recognition, Sight Machinery, financial analysis, product optimization, stock management contracts, management of insurance funds, recognize letters and numbers, diagnosis, and so on (XU et al., 2019). There are several types of Support vector machine, such as integral Support Vector Machine, nonlinear Support Vector Machine vector machine, multi-class Support Vector Machine and fuzzy Support Vector Machine. Most approaches to achieve a system of intelligent behavior are based on components that automatically learn from past experiences. The development of these learning modes is the goal of knowledge known as machine learning. Over the past decade, researchers have made numerous advances in this area by successfully applying machine learning techniques in the tunneling industry (ZHANG et al., 2019). Machine learning has been evolving in recent years and many new specialties are applying these principles in the tunneling industry. Among the various algorithms that exist in the field of machine learning, the support vector machine can be mentioned as one of the most well-known algorithms, used for classification and regression. In the tunneling industry, Ge et al. (2013), calculated TBM performance using least square support vector machine, Mahdevari et al. (2013), developed a model based on SVM algorithm for prediction of tunnel convergence during excavation, Zhang & Gao (2019), studied at SVM regression method in tunnel fires and more. Linear classification methods attempt to separate data by constructing a superstructure (which is a linear equation). Support vector machine classification method, which is one of the linear classification methods, finds the best superstructure that separates the data of two classes with maximum distance for better understanding, Figure (1) shows an image of a two-class dataset that selects the best supercapacitor support vector machine for their separation.

In this section we want to describe how to build a superconducting separator on a detailed example. A detailed illustration of how the superconducting separator is formed by the support vector machine is shown in Figure 2.
Figure 2 - Build a superstructure separator between two data layers in two-dimensional space

First, consider a convex around the points of each class. In Figure 2, it is plotted around points of class -1 and points of class +1 convex shell. The line P is the line that represents the closest distance between two convex shells. h, which is actually the superconducting divider, is a line that halves P and is perpendicular to it. The main idea is to select a suitable separator. I mean a separator that has the greatest distance to the neighboring points of both floors. This answer actually has the largest boundary with points on two different floors and can be bordered by two parallel superstructures crossing at least one of the two floors. These vectors are called backup vectors. The mathematical formula of these two parallel superstructures forming the boundary of the separator is shown:

\[ w \cdot x - b = 1 \]  
\[ w \cdot x - b = -1 \]  

The thing to note is that if the training data are linearly separated, two boundary superpowers can be selected so that no data is between them and then maximize the distance between these two parallel superpowers. Using the geometrical theorems, the distance of these two superstructures is \( \frac{2}{|w|} \). So should \(|w|\) Minimize. The data points within the boundary area should also be avoided, with a mathematical constraint added to the formal definition. For each \( i \), the following constraints are achieved that do not cross the border:

\[ w \cdot x_i - b \geq 1 \]  
\[ w \cdot x_i - b \leq -1 \]

The following limitation can be expressed as:

\[ c_i (w \cdot x_i - b) \geq 1, 1 \leq i \leq n \]  

Solve the issue in general

The problem is to minimize the function \( f(x) \) subject to constraint \( g(x)=0 \). The condition required to \( x_0 \) answer this question is as follows:
\[
\left\{ \frac{\partial}{\partial x} \left( f(x) + \alpha g(x) \right) \right|_{x=x_0} = 0
\]
\[
g(x) = 0
\]  

(6)

For several restrictions \( g_i(x) = 0 \):

\[
\left\{ \frac{\partial}{\partial x} \left( f(x) + \sum_{i=1}^{n} \alpha_i g_i(x) \right) \right|_{x=x_0} = 0
\]
\[
g_i(x) = 0 \quad \text{for } i = 1, \ldots, m
\]  

(7)

If the constraint is unequal, \( g_i(x) \leq 0 \) is still the same problem except that \( \alpha_i \) must be positive. In this case, if \( x_0 \) is the solution to the optimization problem, then for every \( i = 1, \ldots, m \) there must be 0 such that \( x_0 \) applies to:

\[
\left\{ \frac{\partial}{\partial x} \left( f(x) + \sum_{i=1}^{n} \alpha_i g_i(x) \right) \right|_{x=x_0} = 0
\]
\[
g_i(x) \leq 0 \quad \text{for } i = 1, \ldots, m
\]
\[
f(x) + \sum_{i=1}^{n} \alpha_i g_i(x)
\]  

(8)

The function is called Lagrangian. The slope of this function must be set to zero.

**Solve the main issue**

Minimize the issue \( \frac{1}{2} ||w||^2 \) given the limitations 1- \( y_i (w^T x_i + b) \leq 0 \) for \( i = 1, \ldots, n \) In this case, the Lagrangian function is:

\[
l = \frac{1}{2} w^T w + \sum_{i=1}^{n} \alpha_i (1 - y_i (w^T x_i + b))
\]  

(10)

It should be noted that \( ||w||^2 = w^T w \) by derivation \( l \):

\[
w = \sum_{i=1}^{n} \alpha_i y_i x_i
\]  

(11)

\[
w + \sum_{i=1}^{n} \alpha_i (-y_i)x_i = 0
\]  

(12)

\[
\sum_{i=1}^{n} \alpha_i y_i = 0
\]  

(13)

In this study, support vector machine was used to predict penetration rate of TBM. The settings used in SVM to predict the TBM penetration rate are as described in Table 1.
Table 1 - SVM design parameters in this study

<table>
<thead>
<tr>
<th>Model</th>
<th>Kernel</th>
<th>Degree</th>
<th>$\varepsilon$</th>
<th>C</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon$-SVR</td>
<td>Radial Basis Function (RBF)</td>
<td>2</td>
<td>0.1</td>
<td>1000</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Case Study: The Queens Water Tunnel # 3, stage 2

The Queens Water Tunnel # 3, stage 2 was constructed between 1997 and 2000. This section of the tunnel is approximately 7.5 km long and 7.06m in diameter beneath Brooklyn and Queens. The location of The Queens Water Tunnel No. 3, stage 2 as shown in Fig. 3. Descriptive statistic of database in Table 2 (SPSS, 2017).

Table 2 - Descriptive statistic of database

<table>
<thead>
<tr>
<th></th>
<th>UCS (MPa)</th>
<th>BTS(MPa)</th>
<th>PSI (kN/mm)</th>
<th>DPW(m)</th>
<th>Alpha angle (°)</th>
<th>ROP(m/hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>151</td>
<td>151</td>
<td>151</td>
<td>151</td>
<td>151</td>
<td>151</td>
</tr>
<tr>
<td>Minimum</td>
<td>118.3</td>
<td>6.7</td>
<td>25</td>
<td>.05</td>
<td>2</td>
<td>1.27</td>
</tr>
<tr>
<td>Mean</td>
<td>150.053</td>
<td>9.550</td>
<td>34.58</td>
<td>1.0209</td>
<td>44.72</td>
<td>2.0441</td>
</tr>
<tr>
<td>Maximum</td>
<td>199.7</td>
<td>11.4</td>
<td>58</td>
<td>2.00</td>
<td>89</td>
<td>3.07</td>
</tr>
<tr>
<td>Geometric Mean</td>
<td>148.497</td>
<td>9.509</td>
<td>33.71</td>
<td>.7563</td>
<td>36.29</td>
<td>2.0131</td>
</tr>
<tr>
<td>Harmonic Mean</td>
<td>147.021</td>
<td>9.466</td>
<td>32.99</td>
<td>.4526</td>
<td>23.85</td>
<td>1.9822</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>22.1874</td>
<td>.8695</td>
<td>8.462</td>
<td>.64468</td>
<td>23.279</td>
<td>.35982</td>
</tr>
<tr>
<td>Std. Error Mean</td>
<td>1.8056</td>
<td>.0708</td>
<td>.689</td>
<td>.05246</td>
<td>1.894</td>
<td>.02928</td>
</tr>
<tr>
<td>Variance</td>
<td>492.282</td>
<td>.756</td>
<td>71.605</td>
<td>.416</td>
<td>541.912</td>
<td>.129</td>
</tr>
</tbody>
</table>


Figure 3 - The location of The Queens Water Tunnel No. 3, stage 2

MODELING STEPS

Input and output parameters

In this section we call data using MATLAB software (MATLAB, 2018), input and output parameters in Table 3. the database was analyzed through Support Vector Machine.

Table 3 - Input and output parameters

<table>
<thead>
<tr>
<th>Input</th>
<th>Uniaxial Compressive Strength (UCS), Brazilian Tensile Strength (BTS), Peak Slope Index (PSI), Distance between Planes of Weakness (DPW), Alpha angle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output</td>
<td>Rate of Penetration (ROP) (m/hr)</td>
</tr>
</tbody>
</table>

Evaluation criteria ($R^2$ and RMSE)

Coefficient of determination ($R^2$) and Root Mean Square Error (RMSE) are the evaluation criteria in this study which is obtained from the following relationships:

\[
R^2 = \frac{\sum_{i=1}^{N}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{N}(X_i - \bar{X})^2 \sum_{i=1}^{N}(Y_i - \bar{Y})^2}}
\]

\[
RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N}(X_i - Y_i)^2}
\]

$X_i$ and $Y_i$ are the computational and observational values of the time step $i$, $N$ is the number of time steps. $\bar{X}$ and $\bar{Y}$ are the average of computational and observational values, respectively.

Diagram Results

As part of study, the database was analyzed through Support Vector Machine (SVM) modeling. Coefficient of determination ($R^2$), Root Mean Square Error (RMSE), and the Support Vector Machine (SVM) distribution pattern for predicting ROP are shown in Figure 3. This graph represents the fitting line between the values predicted by SVM model with the best fit line $y = x$.

Distribution diagram and fitting diagram of the measured Rate of Penetration (ROP) and the predicted Rate of Penetration (ROP) are shown in Figures 4 and 5, respectively.
Regression analysis

In this step, we analyze the type of regression using Excel software (Excel, 2019) by predicted and measured ROP. As you can see in the figure 6, linear regression is the most appropriate type.
Figure 6 - (a) Exponential; (b) Linear; (c) Logarithmic; (d) Power

(a) Exponential

\[ y = 0.5981e^{0.5925x} \]

\[ R^2 = 0.9453 \]

(b) Linear

\[ y = 1.2267x - 0.4689 \]

\[ R^2 = 0.9678 \]
CONCLUSIONS

The development of underground structures has increased significantly in recent years. Tunnel construction using TBM is an important method used in the tunneling industry. Timing is very important in tunneling projects. A project should be done in a timely manner, otherwise it may have undesirable consequences for the contractors. The time and cost of completing the project is estimated based on the performance of the tunneling machine. Penetration rate is one of the main parameters to check the performance of the machine in the project. There are various methods for predicting this important parameter, each of which has its own characteristics, based on parameters related to rock mass and machine parameters. The purpose
of the present study is to develop a model to estimate TBM penetration rate using support vector machine. Queens tunnel is selected as a case study and the proposed model is evaluated with its data. The coefficient of determination (R²) and Root Mean Square Error (RMSE) in this study were 0.96 and 0.06, respectively. Ultimately, it can be concluded that predictive model lead to acceptable results.

REFERENCES


AFRADI A.; EBRAHIMABADI, A.; HALLAJIAN T. Prediction of the Penetration Rate and Number of Consumed Disc Cutters of Tunnel Boring Machines (TBMs) Using Artificial Neural Network (ANN) and Support Vector Machine (SVM)-Case Study: Beheshtabad Water Conveyance Tunnel in Iran. Asian Journal of Water, Environment and Pollution 16, 1, 49-57, 2019, Available from: https://doi.org/10.3233/AJW190006


AFRADI, A.; EBRAHIMABADI, A.; HALLAJIAN, T.
PREDICTION OF TBM PENETRATION RATE USING SUPPORT VECTOR MACHINE


MICROSOFT CORPORATION. Microsoft Excel, Available at: https://office.microsoft.com/excel, 2019.


AFRADI, A.; EBRABIMABADI, A.; HALLAJIAN, T.
PREDICTION OF TBM PENETRATION RATE USING SUPPORT VECTOR MACHINE

XU, B.; SHEN, S.; SHEN, F.; ZHAO, J. Locally linear SVMs based on boundary anchor points encoding. *Neural Networks*, 117, 274–284, 2019, Available from: [https://doi.org/10.1016/j.neunet.2019.05.023](https://doi.org/10.1016/j.neunet.2019.05.023)


