ORIGINAL PAPER

Predicting Upcoming Collapse Incidents During Tunneling in Rocks with Continuation Length Based on Influence Zone

Sharmin Sarna[1] · Marte Gutierrez[1] · Michael Mooney[1] · Mengqi Zhu[1]

Abstract

This research proposes a model for early prediction of collapse incidents in rocks during tunnel boring machine (TBM) operation utilizing TBM and geological survey data. Uncertainties associated with tunnel geology can significantly hamper tunneling progress. So far, various research works have tried to reduce such uncertainties by predicting the geological condition during tunnel excavation. However, specific tunneling responses caused by adverse geology are still poorly studied. In rocks, adverse geology such as karstic caves and fault zones can cause tunnel boundary collapse, a detrimental tunneling response causing unexpected interruptions and casualties. To identify and predict the potential collapse of TBM tunnels in rocks, three machine learning (ML) classifiers are used, namely: (1) multilayer perceptron, (2) support vector machine, and (3) random forest. The ML algorithms are trained and validated using data on collapse incidents in a water conveyance tunneling project in China. The prediction accuracy of the proposed model reached 98% for training data and 97% for validation data. Furthermore, the model can identify an "Influence Zone" for a collapse incident. A unique contribution of this research is that the "Influence Zone" enables the model to predict an impending tunnel collapse and the extent of that collapsed segment ahead of the excavation. Finally, to gain better insight into the ML-based predictions, the relationships between TBM-related features and tunnel geology are carefully analyzed.

Highlights

- Machine learning classifiers can detect a zone of influence for approaching ground collapse during tunnel excavations through disturbed rock.
- The zone of influence is found to be proportional to the continuation length of the collapsing ground.
- Cutterhead rotating speed of tunnel boring machine and rock mass ranks are the two most important indicators for collapsing ground conditions.

1 Introduction

Geological uncertainties experienced during a tunneling project can pose severe threats to the construction progress by causing operational delay, casualties, cost overrun, etc. (Delisio and Zhao 2014; Li et al. 2017; Macias et al. 2014;

✉ Sharmin Sarna
  sarna@mines.edu

[1] Department of Civil and Environmental Engineering, Colorado School of Mines, Golden, CO, USA

Zhu et al. 2021). There is a financial limitation for conducting geological/geotechnical tests for a tunneling project. Therefore, the geological survey data collected before the tunnel construction are often insufficient when a tunnel boring machine operates through an adverse geologic condition. At the same time, geological characterization techniques such as seismic reflection, 3D polarization, resistivity method, electromagnetic method, etc. are often installed to get early warning on possible detrimental conditions ahead of the TBM (Alimoradi et al. 2008; Li et al. 2017, 2018; Mooney et al. 2012; Wang et al. 2019a, b). However, these monitoring systems require additional time, costing, and

maintenance for the tunneling operation. Therefore, recent research focused on using operational parameters recorded by the TBM data acquisition system to predict the geologic condition ahead of the tunnel face (Exadaktylos et al. 2008; Jung et al. 2019; Liu et al. 2019; Yamamoto et al. 2003).

TBM data acquisition system includes multiple sensors to collect data with high frequency during the tunneling operation (Girmscheid and Schexnayder 2003; Moreno et al. 2015). Various sensors provide an extensive operational database suitable for performing machine learning and statistical analysis (Sheil et al. 2020). These data represent the interaction of the TBM with the excavated ground. Previous research works attempted to utilize the TBM collected data and its interaction with the surrounding ground to ensure better TBM operation and performance (Benato and Oreste 2015; Chen et al. 2017; Gao et al. 2019; Gong et al. 2021; Gong and Zhao 2009; Hassanpour et al. 2011; Khademi Hamidi et al. 2010; Salimi et al. 2016, 2017; Sapigni et al. 2002; Sun et al. 2018; Yin et al. 2014).

Several researchers have developed prediction models based on statistical and machine learning methods to deal with the uncertainties of the ground conditions and response by utilizing the TBM operational data. Jung et al. (2019) used artificial neural network (ANN) with Levenberg–Marquardt (L–M) minimization of back-propagation error to predict mixed face ground condition one ring ahead of the tunnel face using TBM data on torque (T), thrust (F), and penetration rate (P). The ANN algorithm used data from different tunneling projects with ground conditions varying from soft soil to hard rock. Zhang et al. (2019) used TBM operational data of four channels, namely, torque (T), thrust (F), cutterhead rotation (R), and advance rate (P), to predict rock mass type of the surrounding ground with ML classification algorithms (support vector machine, k nearest neighbor, random forest). Liu et al. (2020) used an ensemble learning model based on classification and regression tree and the AdaBoost algorithm (Géron 2019) for real-time determination of Chinese hydropower class (HC) of the surrounding rock using TBM operational and performance parameters. Mito et al. (2003) used drill logging and TBM operational data to assess the geological conditions of the surrounding ground ahead of the tunnel face using geostatistical techniques.

Despite the attempts made so far, detailed studies are still required to characterize different tunneling responses caused by various adverse geologic conditions in rock strata. Potential adverse geologies encountered during tunneling operation in rock strata include karst caves, faults or fractured zones, rockburst, blocky rock, etc. (Jeong et al. 2018; Macias et al. 2014; Rostami 2016). Rock tunneling in karst or fault fractured zones can cause detrimental tunneling responses, such as TBM blockage, tunnel boundary collapse, water inrush, etc. (Huang et al. 2018; Parise et al. 2008).

Tunnel boundary collapse is defined as the disintegration of collapsing blocks from the excavated tunnel boundary due to the presence of weak or unstable rock caused by fault zones or karstic intrusions encountered along the tunnel alignment (Fraldi and Guarracino 2009, 2010; Wang et al. 2019a, b; Wang et al. 2017; Yang et al. 2017). Several research methods have been made to understand the collapse mechanism in disturbed rock conditions and predict them during tunnel operation. Huang et al. (2020) utilized an analytical approach to predict collapse region during a deep highway tunnel excavation describing the conditions inducive of rock mass collapse having a karstic cave above the excavated tunnel boundary. An upper bound theorem of limit analysis was utilized to develop an analytical expression of the surface of the collapsing block near the area of the karstic intrusion. This method depends on the knowledge of the exact location of the karstic cave concerning the tunnel excavation boundary, the local rock mass properties, such as compressive and tensile strength parameters, unit weight, material constants for the Hoek and Brown failure criterion, etc. The approach utilized available rock mass properties local to the vicinity of the karstic cave to represent the actual rock mass condition of that region. However, abundant data on rock mass properties would require frequent geological and geotechnical testing and investigation along the tunnel alignment, which is impractical for tunneling projects.

In addition, information gathered by geophysical sensors is needed to prospect karstic cave locations along the tunnel alignment, which incurs additional time and cost. Moreover, this method mainly focuses on investigating the collapse mechanism's nature during the excavation rather than predicting collapses in advance during run-time. The latter could potentially benefit the tunneling crew in taking proper countermeasures in advance, thus avoiding unnecessary delay and economic losses. Xue et al. (2020) utilized an analytical hierarchy-process entropy method and fuzzy set theory to establish collapse risk indicators for soft rock tunnels based on eight selected influencing factors.

The evaluation model developed can evaluate the collapse risk grade defined by the range of deformation of the support structure during tunnel construction in soft rock conditions. However, this analysis method is also dependent on the geological and geophysical survey, and testing data as this method can only predict the risk grades of a tunnel section when these data are available. Moreover, the model is established and validated utilizing very few data points, making it more susceptible to outliers or exceptional conditions during tunnel construction. On the other hand, a data-driven prediction model established and validated on real-time data collected by the TBM sensors at a high frequency can be more robust and reliable. Chen et al. (2021) adopted such an approach, where a time-series forecasting method was combined with a deep belief network

to establish a neural network prediction model predicting the value of a parameter named as drilling efficiency index calculated utilizing TBM operational parameters. The predicted index can identify collapsing ground based on its deviation from the actual value. However, in a practical scenario, forecasting an upcoming collapse incident ahead of the existing tunnel face and its entire length will be more beneficial for the tunnel operator to take necessary precautions and preventative measures. Guo et al. (2022) developed a three-stage method for forecasting tunnel ground collapse during TBM excavation by training deep learning algorithms and LSTM. The developed method can successfully forecast collapse ground condition based on anomalies in accuracy measure of predicted torque and thrust by the LSTM method and associated rock grade. Nevertheless, as opposed to a time series forecasting method, such as LSTM, a simple classification method is less time-consuming and independent of immediate previous measurements to make a prediction. Therefore, this research aims to forecast collapse incidents during TBM tunneling in adverse rock conditions while providing a simple method to estimate the intensity of that upcoming collapse. The study used data from a water conveyance tunneling project with several collapse incidents, and thus developed prediction models for the early prediction of tunnel boundary collapses by training state-of-the-art machine learning classifiers. The proposed model can predict an upcoming collapse incident during tunneling and the extent of the collapsing ground ahead of the tunnel excavation. Considering the level of uncertainties the proposed model can resolve, it is hoped that significant progress can be made for TBM tunneling through difficult ground conditions with the potential of frequent collapse incidents.

Following the description of the tunneling project used in this study, an "Influence Zone" for collapses, which provides new insight regarding ground conditions that lead to collapse, is described. Data pre-processing techniques are used to remove spurious data points, ensuring optimal utilization of TBM data for the input to the ML model. Three ML classifiers are trained to build prediction models consisting of multilayer perceptron, support vector machine, and random forest. Model predictions are tested and validated against the field data. Also, multiple ML models are trained and tested to verify and propose the optimal length of the "Influence Zone". Finally, a detailed analysis of the TBM operational data is conducted to gain better insight into their variations in response to tunnel geology to reduce the black-box nature of the ML-based models.

2 Tunnel Project Overview and Data Description

This section provides an overview of the tunneling project and a detailed description of the geological and the TBM data used in the study.

2.1 Geology

This study utilized data from a water conveyance tunneling project near Jilin Province, China. This project has a total tunnel length of 69.89 km for the first three sections, starting from the Fengman reservoir near Songhua River and running until the Shuangyang reservoir near the Yinma river. The tunnel was excavated by simultaneous TBM and drill-and-blast (D&B) methods. The D&B method was applied, where TBM could not operate due to disturbed ground conditions. This study collected data from the third tunneling section, representing the most challenging construction segment due to complex geologic conditions. This tunnel segment has a total length of about 20 km (from chainage $71 + 476$ m to $51 + 705$ m), in which 88% of excavation was done with an open mode gripper TBM manufactured by the China Railway Engineering Equipment Group Co. Ltd (CREC). The rest of the excavation was performed with the D&B method.

The primary geology in this segment includes limestone, granite, tuff, and diorite. Among these rock types, limestone and granite are the most dominant lithologies. The detailed description of the excavation method and lithology along the tunneling alignment is shown in Table 1. The average overburden depth at this tunnel section is about 100 m. Figure 1 shows the longitudinal geologic profile of this tunnel section. The tunneling operation of this section continued for 803 days, including 76 days of construction shutdown.

The geological profile was characterized and recorded on-site during the excavation process (Jing et al. 2019). The surrounding rock mass was classified based on the HC (Chinese Hydropower Class) System, which divides rock masses into five classes, I to V (Liu et al. 2017). Class I represents the strongest rock mass in the HC system, and Class V the weakest. Several rock characteristics are correlated with this classification system, including the uniaxial compressive strength, fracture intensity, discontinuity, groundwater condition, etc. In the present study, the surrounding rock mass of the tunnel includes Class II to Class V, among which Class III is the dominant rock mass type (Table 2). The rock mass uniaxial compressive strength varies from 38 to 95 MPa, and

Table 1 Information about lithology of tunnel alignment

Lithology	Chainage (m)		Total length (m)	Percentage (%)	
	Start	End		TBM	D&B
Limestone	$71 + 476$	$63 + 266$	8210	39	2
Diorite	$63 + 266$	$62 + 978$	288	1	0
Tuff	$62 + 978$	$58 + 254$	4724	15	10
Granite	$58 + 254$	$51 + 705$	6459	33	0
Summary			19,681	88	12

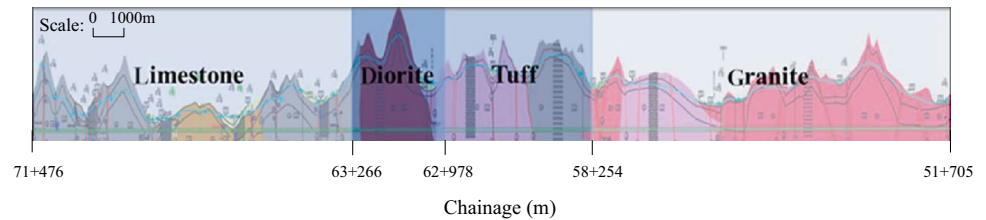**Fig. 1** Longitudinal geologic profile of tunnel Sect. 3 (Chen et al. 2021)

Chainage (m)

Table 2 Rock mass classification according to the HC system with percentages

Rock mass class	Percentage of tunnel length (%)
II	5.6
III	75.1
IV	15.4
V	3.9

Table 3 Information on collapse locations

Lithology	No	Start Chainage (m)	End Chainage (m)	Collapse length (m)
Limestone	1	70+817	70+800	17
	2	70+792	70+768	24
	3	70+396	70+380	16
	4	70+348	70+335	13
	5	66+329	66+307	22
	6	66+290	66+274	16
	7	66+243	66+217	26
Tuff	8	60+334	60+328	6
Granite	9	56+602	56+597	5
	10	56+574	56+556	18
	11	56+495	56+474	21
	12	56+313	56+303	10
	13	56+291	56+281	10
	14	56+241	56+231	10
	15	55+003	54+963	40
	16	54+954	54+944	10
	17	54+900	54+847	53
	18	54+216	54+170	46

Table 4 TBM specifications

Parameter	Value
Type	Open Mode Gripper TBM
Excavation diameter	8.03 m
Cutterhead diameter	7.93 m
Nominal disc-cutter spacing	89 mm
Cutterhead driving power	3500 kW
Maximum thrust of cutterhead	23,260 kN
Nominal torque of cutterhead	8410kNm @ 3.97 rpm
Maximum rotational speed of cutterhead	7.6 rpm
Maximum advancing speed	120 mm/min
TBM weight	180 t
Thrust cylinder stroke	1800 mm
Number of cutters on cutterhead	56 (48 face cutters (19 in), 8 edge cutters (17 in)

2.2 TBM Description

The TBM used in this project is an open mode gripper TBM with a 7.93 m face diameter and 8.03 m excavation diameter. A detailed description of the TBM specifications is provided in Table 4.

The data-driven approach of modern-day TBMs relies on a central data acquisition and monitoring system controlled by a programmable logic controller (PLC) (Gong et al. 2021; Mooney et al. 2012). This data acquisition system collects real-time data on various TBM operational parameters, such as cutterhead rotation, torque, axial force, and displacement of the thrust cylinders, which are eventually used to calculate the total thrust force and an advance rate of the TBM. Figure 2 shows a schematic diagram of an open mode gripper TBM with some sensor locations. These sensors are connected to the PLC system for collecting the operational parameters mentioned above. The PLC system can transfer these data through a communication interface provided by the equipment manufacturer. Hence, the parameters are accessible to the monitors of the TBM control cabin, where the operator can visually inspect the data to check against any unusual activity and initiate necessary action. These data are also stored in real-time in local and/or remote databases. For the current project, the TBM data acquisition system
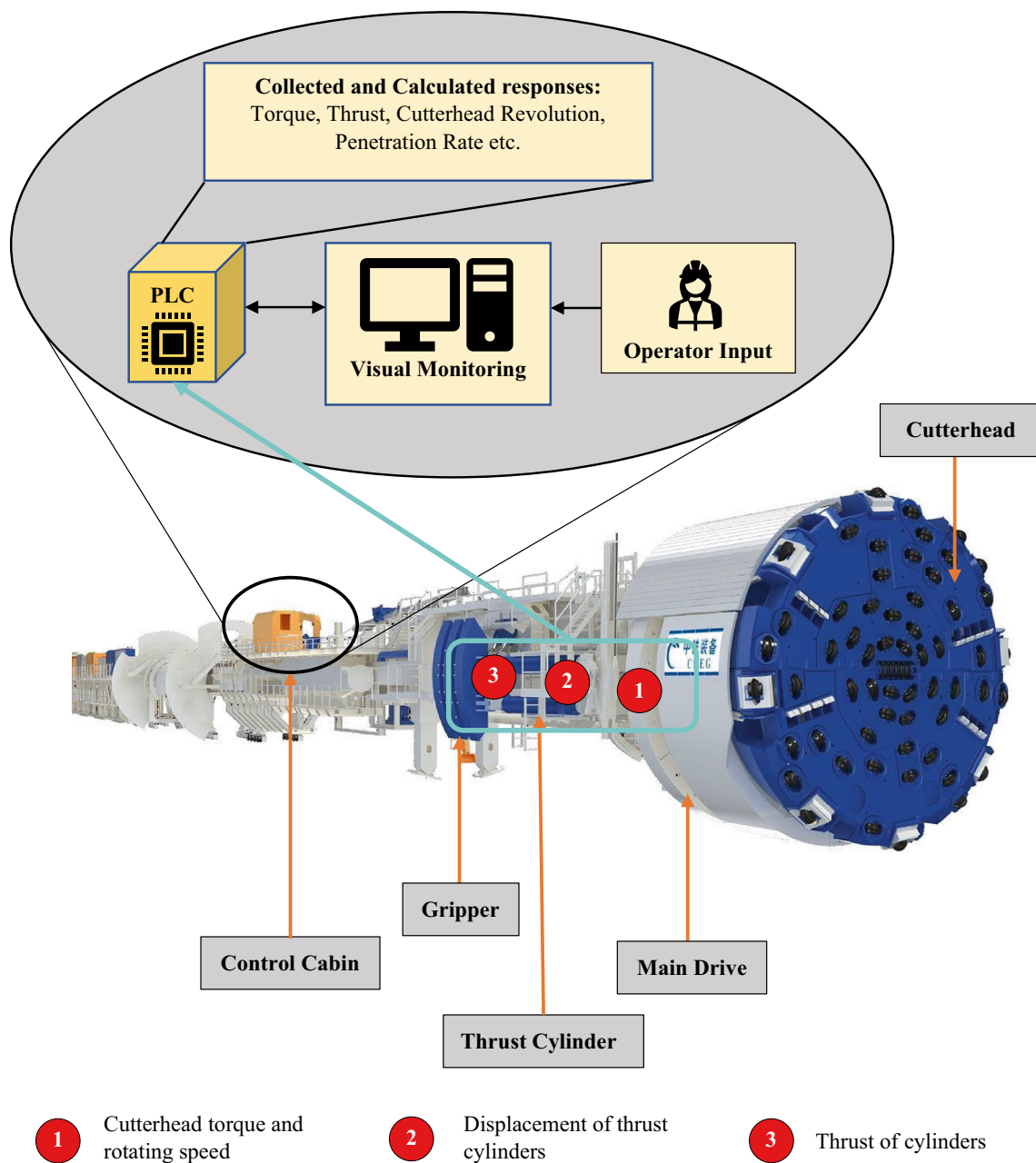
the volumetric joint count ranges from 3.8 to 25.68 per cubic meter (Zhu et al. 2021).

The tunneling project suffered frequent collapses caused by faulted and fractured zones and karstic caves. Most of these collapses are located at the tail of the TBM shield (Chen et al. 2021). There are 18 records of such locations, shown in Table 3.

Collected and Calculated responses:
Torque, Thrust, Cutterhead Revolution,
Penetration Rate etc.

PLC

Visual Monitoring

Operator Input

Cutterhead

Gripper

Control Cabin

Main Drive

Thrust Cylinder

1 Cutterhead torque and rotating speed

2 Displacement of thrust cylinders

3 Thrust of cylinders

Fig. 2 Schematic diagram of gripper TBM with sensors' arrangement

collected data at a 1 Hz frequency. Figure 3 displays the raw TBM data as visualized in the control cabin.

The TBM excavation was performed in tunneling cycles. Each cycle is expected to excavate approximately equal to the TBM thrust cylinder stroke, which is 1.8 m. However, the surrounding ground condition affects the TBM performance, resulting in different tunneling cycle lengths (varying from 0.5 to 1.8 m). During excavation, the TBM data acquisition system collected data automatically from 199 sensors with a 1 Hz frequency, which resulted in a data

volume of terabytes. These 199 sensors primarily represent TBM operational parameters.

A typical tunneling cycle includes four segments: (1) shutdown, (2) free rotation, (3) ascending, and (4) steady operation. Initially, the TBM starts with the free rotating segment, with the TBM rotating freely before hitting the ground. When the TBM hits the ground, the operational parameters gradually increase in the ascending or rising segment. Ultimately, the TBM excavation operation starts to stabilize as the fluctuation of the operational parameters reduces, which is regarded as the stable segment. The
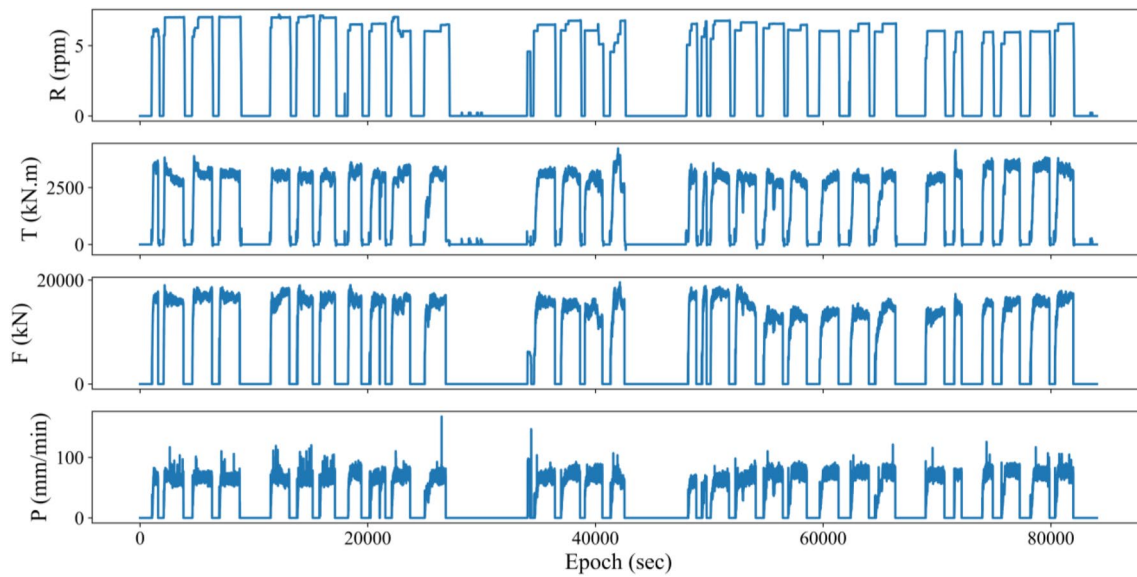
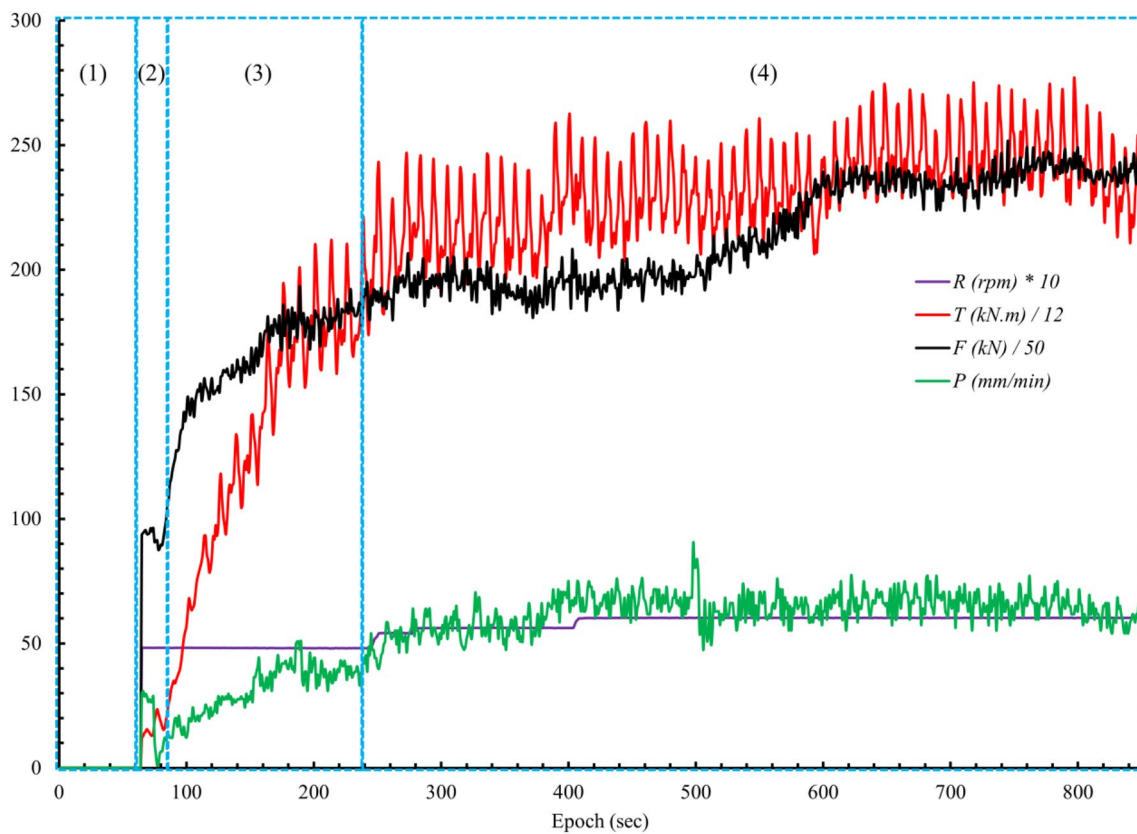**Fig. 3** Examples of raw TBM data

Fig. 4 Typical tunneling cycle showing four TBM operational parameters with time at different stages: (1) shutdown, (2) free rotating, (3) ascending, (4) stable operation segment

TBM operates in the stable mode for a while performing the main excavation operation. Finally, the TBM stops for some time, called the shutdown segment, before starting the next tunneling cycle. A typical cycle is shown in Fig. 4 with four segments for the above-mentioned operational parameters.

3 Influence Zone for Tunnel Collapse

In the case of a homogeneous rock mass, the TBM excavation process continues regularly without significant deviation in the responses of the TBM. However, when excavating through an adverse geologic condition, the TBM operates irregularly with frequent interruptions, such as machine blockage or water inrush, requiring unplanned machine shutdown and repair. These situations are unwanted for the TBM operator. In addition, the TBM operator has limited knowledge of the ground conditions that the TBM is excavating, as it is very challenging to predict the ground condition ahead of the TBM location. Therefore, an early warning for excavating through such ground conditions can help the construction team take better precautions. Hence, understanding the nature of the ground, while the TBM is operating is beneficial.

The nature of the ground surrounding the adverse geologic condition can help the TBM operator anticipate the proximity of a potential difficult tunnel section. Any disturbance in a homogeneous ground condition will have some boundary effects on its surrounding, affecting the ground resistance experienced during tunnel excavation. Hence, it is assumed that the ground surrounding the collapse locations (representing adverse geologic conditions, e.g., karstic cave, fault fracture zone) will behave differently than the normal ground condition, while TBM excavation continues. These ground locations are assumed as "Influence Zone." This study focuses on identifying these locations from the normal ground conditions during TBM operation. Successful identification of these locations can warn the TBM driver of any approaching collapse ground while excavating.

The tunneling project utilized in this research has 18 different collapse locations with varying lengths of collapse experienced while operating the TBM (Table 3). It is commonly understood that these collapses are different in their level of influence on the adjacent ground. A longer collapse is expected to have a longer "Influence Zone" than a shorter length collapse. Hence, this study assumes two "Influence Zone" locations for each "Collapse Zone" location. One location is before the "Collapse Zone," and another is after the "Collapse Zone." The lengths of these "Influence Zone" locations are decided according to the lengths of their corresponding "Collapse Zone," as shown in Fig. 5. For each "Collapse Zone" with length L, two "Influence Zones" having $L/4$ length each are assumed. The rest of the tunneling alignment is considered a "Normal Zone." Later, machine learning classifiers are trained to categorize these three types of ground conditions.

4 Data Pre-processing

This section describes the pre-processing method of raw TBM data and features used for ML prediction models.

4.1 TBM Data

The raw TBM data collected for each tunneling cycle include the shutdown, free rotation, ascending, and stable segments (Fig. 2). Among these segments, the stable segment represents the main excavation operation. Hence, only the stable segments' data are collected from the tunneling cycles during the pre-processing of raw TBM data. Another source of data impurity is outliers. The outliers are sourced by the sudden jump of the sensors collected data due to sudden machine malfunction or other unforeseen reasons.
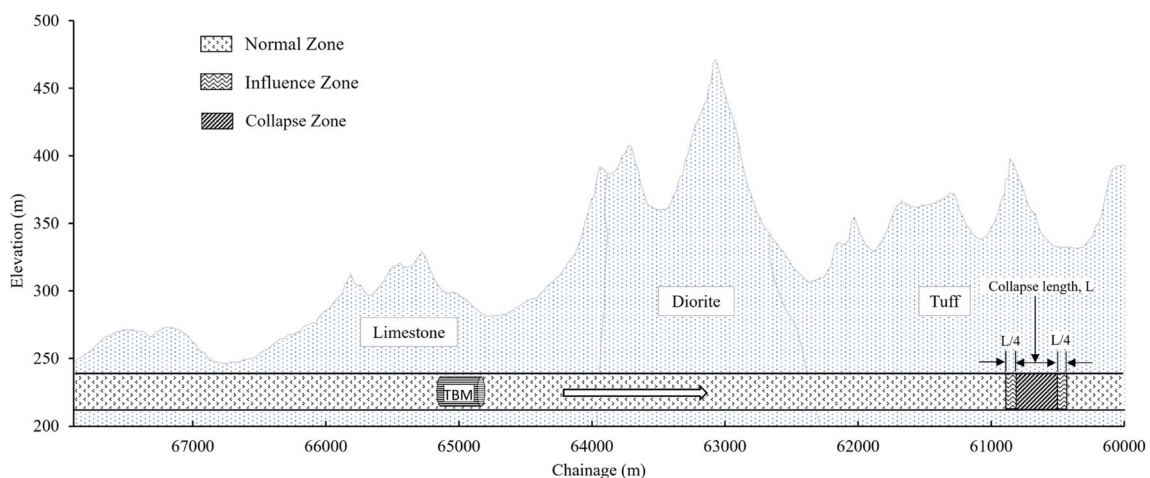

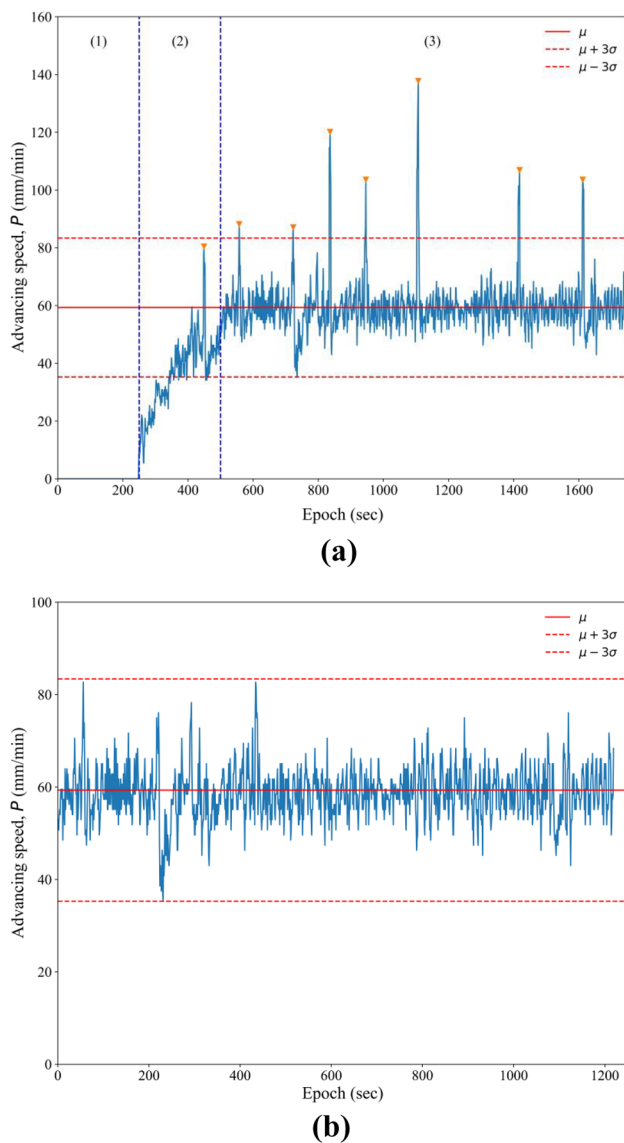**Fig. 5** "Influence Zone" before and after "Collapse Zone"

Fig. 6 Pre-processing of TBM data: (**a**) raw TBM data; (**b**) stable TBM data after removing shutdown (a (1)), free rotating and ascending segments (a (2)), and outliers with three-sigma truncations

They are not representative of the surrounding condition and deserve to be removed from the database. This task applies the "three-sigma truncation rule" to the stable segment data. This method removes all the data farther apart from the three standard deviations of the mean of stable segment data of a tunneling cycle. These processes are illustrated in Fig. 6 for the advancing speed, P (mm/min) of a typical tunneling cycle.

The TBM sensors' collected data are preserved in original form after the above-mentioned pre-processing mechanism to avoid loss of information. Compressing the TBM data over the tunneling cycles (getting a single value, e.g., mean of a parameter for each cycle) is less representative of the

ground condition along the tunnel alignment. Data compression is also less beneficial when the focus is required for specific locations which are scarce in number. In this study, very few collapse locations are recorded, comprising 363 m of tunneling operation. In this particular case, keeping the original TBM data without any compression benefits two ways: (1) sufficient data points are achieved representing the collapse locations, (2) optimal utilization of the high-frequency data collection practice of the TBM data acquisition system is ensured.

4.2 Feature Selection and Description

As mentioned in Sect. 1, several researchers (Jung et al. 2019; Liu et el. 2019; Zhang et al. 2019; Liu et al. 2020, Chen et al. 2021) have successfully utilized four TBM operational parameters to detect the surrounding ground condition of the TBM, namely, cutterhead torque, T (kN-m), cutterhead rotating speed, R (revolution per minute or rpm), advancing speed, P (mm/min), and thrust or propulsion, F (kN), as well as three TBM performance parameters (Liu et al. 2020; Yamamoto et al. 2003). Among them, R and P are driver operating parameters (Guo et al. 2021). There are three major factors that come into play while TBM driving: the tunnel geology, the human decision, and the machine response. The R and P parameters are affected by all three of them, whereas T and F can represent machine–ground interaction only. Although, the human factor can't be removed from R and P, they still include valuable information about the ground condition due to their sense of machine–ground interaction. Hence, in this study, these parameters are considered as adequate to detect the collapse ground condition and thus selected as input features for the ML prediction models. The three performance parameters are named as the Field Penetration Index (*FPI*) (kN.rev/mm), Torque Penetration Index (*TPI*) (kN.rev), and Specific Energy (E_S) (kN.m/m^3 or kJ/m^3), which are defined as follows:

$$FPI = \frac{F \times R}{no.\,of\,cutters \times P} \tag{1}$$

$$TPI = \frac{T \times R}{no.\,of\,cutters \times P} \tag{2}$$

$$E_S = \frac{T \times 2\pi R \times 10^3}{A \times P} + \frac{F}{A} \tag{3}$$

where A is the cross-sectional area of the TBM cutting face; *FPI* and *TPI* indicate the thrust (F) per cutter and torque (T) per cutter, respectively, for unit penetration per revolution; and E_S indicates the energy required to excavate a unit volume (m^3) of the rock mass. Besides the TBM-related
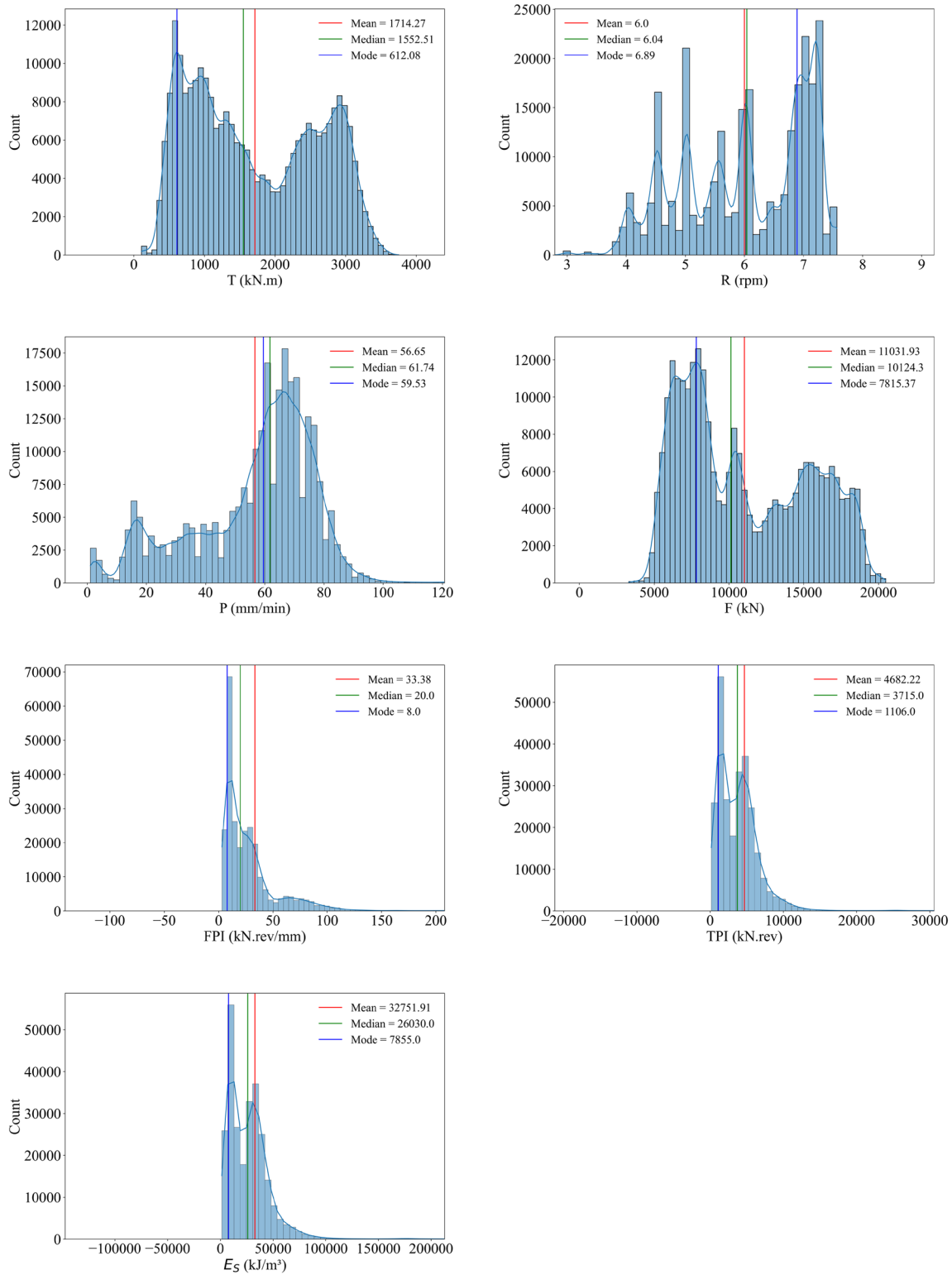
Fig. 7 Histogram plots with data statistics of the TBM-related features

features, Rock Mass Rank (*RMr*) by the HC system is also incorporated as a geological feature variable for the ML prediction models. To mention, the *RMr* values were assigned

to uniform geological sections along the tunnel alignment using on-site logging data during TBM excavation. This feature indicates rock mass quality and is supposed to have a

high potential in detecting collapse conditions of the ground. In total, eight input feature variables are considered for the classification task. Figure 7 shows the histogram plots of the TBM-related features for three standard deviations on either side of the mean, along with density distribution curves and some data statistics, which provide insight into the variability and distribution of these features. Based on the existing literature and expert judgment, the selected parameters were considered sufficient to predict the collapse ground condition. Therefore, no additional feature selection method was applied to the TBM database. Also, the number of features utilized did not challenge the computational or storage capacity, hence considered without any dimensionality reduction to the data set.

5 Machine Learning Classifiers to Build Prediction Model

5.1 Data Set

As mentioned in Sect. 3 of this paper, this study aims to distinguish between the "Collapse Zone" and "Influence Zone" of the tunnel alignment from the "Normal Zone," which will benefit the TBM pilot to mitigate the upcoming hazards by taking early informed decisions during excavation operation. The data set used in this study includes nine feature variables in which eight are input feature variables (*T, R, P, F, FPI, TPI, E_S* and *RMr*). One is output/target feature variable named as 'label.' Among the input features, seven (*T, R, P, F, FPI, TPI, E_S*) are related to the operational parameters collected by the TBM data acquisition system. The *RMr* is obtained by on-site geological survey. The input features are continuous variables (except *RMr*), and the output feature is a categorical variable. To mention, *RMr* values are assigned to representative tunnel sections (Liu et al. 2017), whereas the TBM data used in this study are in the form of direct sensor measurements in seconds. Hence, all the TBM data that are collected within the range of chainage location of such a tunnel section are considered to have that section's *RMr*.

There are three labels or classes, namely, 'Normal,' 'Collapse,' and 'Influence Zone' under the target feature. These classes are assigned to each data instance according to their chainage location (Fig. 5). In this study, machine learning classification algorithms are trained, where the data instances are independent of each other. However, the problem design expects the predicted classes to come sequentially as 'Normal' ground, 'Influence Zone', and 'Collapse' ground condition in the field during tunneling operation while approaching a collapsing ground. Hence, the prediction to be made by the input parameters collected at $(n+1)^{th}$ second is technically independent from

the prediction made by the input parameters collected at n^{th} second. However, the predicted classes are expected to follow the spatial sequence mentioned above while tunneling, as the labels are assigned by the problem design following the field experience. For the 'Normal' class data, only a portion of the total available data is considered to avoid extreme data imbalance in the data set (as most of the tunneling alignment is represented by the "Normal Zone"). Hence, we collected a representative portion of the available 'Normal' class data by maintaining the overall ratio of four *RMr* (Class II–Class V) ranks available along the whole tunneling alignment (Table 2). Hence, the data set is a $9 \times 264,666$ matrix with nine columns and 264,666 rows or instances. The data are distributed as follows: 'Normal' instances: 104,657; 'Collapse' instances: 77,706; 'Influence Zone' instances: 82,303). This data set is used to train and test three ML classifiers for the multiclass classification task. The training data set is prepared by a random split of 70% of the total data set, and the rest of the 30% is preserved as a test data set (train set: $9 \times 185,266$; test set: $9 \times 79,400$).

5.2 Machine Learning Classifiers

The purpose of an ML classifier is to classify data instances into a finite set of categories (Shalev-Shwartz and Ben-David 2014). In this research, we adopted three different ML classifiers, namely: support vector machine (SVM), multilayer perceptron (MLP), and random forest (RF), for classifying the labels of the target feature. MLP and RF are inherently multiclass classifiers, whereas SVM is a strictly binary classifier (One vs. One) (Géron 2019; Pedregosa et al. 2011). These classifiers are chosen for their robustness and outstanding performance for classification tasks. They are described briefly in the following sub-sections.

5.2.1 Multilayer Perceptron Classifier (MLP)

The multilayer perceptron is a neural network-based classification algorithm. Artificial neural networks can handle complex nonlinear data patterns and trends (Géron 2019). It is shown that a two-layer back-propagation neural network with sufficient hidden neurons is a universal approximator (Hornik et al. 1989). The network of an MLP Classifier can be described as a structure of multiple layers of neurons. The connections between the neurons of the layers pass the output of a neuron of the previous layer to the neuron's input to the next layer (Shalev-Shwartz and Ben-David 2014). Initially, arbitrary weights are assigned to all the connections for each neuron of a hidden layer to get a weighted sum of input features. This weighted sum is then passed through an activation function (e.g., *logistic*, *ReLU*, *tanh*, etc.) (Géron 2019) and used as the input for the neurons of the next layer.

Finally, the output layer provides an output compared to the target output to estimate error. The error gradients of multiple iterations are then backpropagated to the previous layers to update the weights using gradient descent to reduce the errors. The activation function allows the gradient descent to get derivatives and progress with each iteration. For classification tasks between multiple (more than two) exclusive classes, the activation function used for the output layer is soft-max or multinomial logistic (Géron 2019).

5.2.2 Support Vector Machine (SVM)

Support vector machine is a very powerful and versatile classification algorithm capable of performing nonlinear classification (Géron 2019). It is suitable for a small to medium size data set with high dimensionality. SVM has the advantage of increasing class separation, which reduces the expected prediction error (Xia 2020). The SVM classifier aims to identify decision boundaries between classes that satisfy the "maximum margin." This margin is defined by the distance of the boundaries from the closest training instances called the "support vectors." A "hard margin" is applied in linearly separable classes without allowing any data instance inside the margin. However, a "soft margin," more flexible and less sensitive towards outliers, is applied in a nonlinear data set. The application of "soft margin" is performed using a regularization hyperparameter C. In a nonlinear data set, the margin optimization is performed in a higher dimensional space instead of the original feature space by applying basis functions to get a linear separating hyperplane. The process is performed by initially using basis functions to the feature space and then optimizing, called the "kernel trick," to avoid computational complexity. Thus, SVM can adopt different kernels (*linear*, *ploy*, radial basis function or *rbf*, etc.) (Géron 2019) to find the best margin and prediction for a classification problem (Kelleher et al. 2020; Wu et al. 2008).

5.2.3 Random Forest (RF)

Random forest is an ensemble learning algorithm built with multiple decision trees. The main benefits of using RF classifier are: (i) very fast with a large data set of high dimensionality, (ii) robust to multicollinearity, noises, and outliers, (iii) less likely to overfit on training data set (better generalization than a single decision tree), (iv) very high accuracy (Breiman 2001; Zhu et al. 2021). RF collects predictions from multiple decision trees trained with different random sample subsets prepared via bagging and random feature subsets. These predictions are averaged, and finally, the class with the highest average probability score is assigned (Ellis et al. 2014). CART (Classification and Regression Tree) algorithm is used to optimize the cost function of each decision tree, which recursively runs till the maximum depth or purity condition is reached. RF searches for the best feature among a random subset of features instead of searching for the very best feature that increases the trees' randomness and diversity. This provides a lower variance for a higher bias, thus yielding an overall better prediction model (Géron 2019).

While training the ML classifiers, the grid search method was adopted to select the best hyperparameter combination for each algorithm. The grid search method performs a search for the hyperparameter values for the best prediction on the trained data within a defined hyperparameter space (Pedregosa et al. 2011). The hyperparameter tuning can also address the overfitting issue in the model, if any, so that the model can perform well on both the "train set" and "test set" data. For the MLP classifier, the combination of *hidden layers*, *nodes*, *activation function*, and *solver* is considered for hyperparameter tuning, which is essential to decide the classifier's performance. For SVM, the regularization hyperparameter C, *kernel*, γ (for *rbf* kernel) are considered for grid search tuning. Finally, for RF, three important hyperparameters, namely: *n_estimators, max_depth, and max_feature* are tuned through grid search (Table 5).

5.3 ML Classifier Performance Measure

As mentioned in Sect. 5.1, the data set is split into a "train set" and a "test set." The ML classifiers are initially fitted on the "train set" by following the tenfold cross-validation method. Finally, the trained models made predictions on the "test set" to validate their performance. Thus, the models are initially trained on tenfolds or disjoint sets of equal size collected from the "train set" using a bootstrap method and utilizing nine of them to train and one to test. The process is repeated for all the ten folds or sets. The performance of a model trained by tenfold cross-validation is an average result of all the tenfolds representing the model's performance

Table 5 Hyperparameter selection through grid search

Classifier	Parameters	Optional values	Best Parameter
MLP	*Hidden layer size*	1,2,3,4	3
	Activation	logistic, ReLU, tanh	ReLU
	Solver	lbfgs, sgd, adam	adam
	Kernel	rbf*, poly, linear	rbf
SVM	C	0.5, 1, 10	1
	γ	0.01, 0.1, 1, 5	5
	n_estimators	25, 50, 75, 100	50
RF	*max_depth*	5, 10, 15, 20	15
	max_features	2, 4, 6	4

*rbf = radial basis function

Table 6 Confusion matrix for a three class-classification

True class	Predicted class		
	Influence zone	Collapse	Normal
Influence zone	+ ve	− ve	− ve
Collapse	− ve	+ ve	− ve
Normal	− ve	− ve	+ ve

* + ve = True predictions, -ve = False predictions

on the overall "train set." Finally, these trained models are tested on the "test set" to perform their validation test.

The prediction results of the classifiers are shown in confusion matrices. Three classes are predicted, so each confusion matrix has three rows and three columns. The total count for each row indicates the "True class" of each of the three classes. The columns indicate the "Predicted class" by the ML classifiers. Thus, each matrix cell shows the count for its "True class" and "Predicted class." All the "True predictions" ("True class" = "Predicted class") are along the diagonal of the matrix, whereas all the "False predictions" ("True class" ≠ "Predicted class") are shown off-diagonal. The detail is shown in Table 6.

The performances of the ML classifiers are measured in terms of performance metrics. The most popular performance metrics for ML classifiers are adopted, namely, *Accuracy, Precision, Recall,* and *F-1 score* (Giussani 2020; Shalev-Shwartz and Ben-David 2014; Sokolova and Lapalme 2009). For a specific class, four types of predictions are possible. They are true positive (*tp*), true negative (*tn*), false positive (*fp*), and false negative (*fn*). Parameters *tp* and *tn* are all the true/correct predictions, and *fp* and *fn* are all the false/incorrect predictions by a classifier. Let us consider the class 'Normal.' If the classifier predicts a 'Normal' class as 'Normal,' it is a *tp* prediction for that class. If a 'Collapse' class appears and the classifier can identify that it is not a 'Normal' class (predicts either 'Collapse' or 'Influence Zone'), it is a *tn* prediction. However, if the classifier predicts a 'Collapse' or 'Influence Zone' class as a 'Normal' class, it is a *fp* prediction. Finally, if the classifier

fails to classify a 'Normal' class as such (i.e., predicts either 'Collapse' or 'Influence Zone'), it is an *fn* prediction.

The performance metrics are defined based on these prediction counts. If there are *c* number of classes to predict, the metrics can be expressed as follows:

$$Accuracy = \frac{\sum_{i=1}^{c} \frac{tp_i + tn_i}{tp_i + fn_i + fp_i + tn_i}}{c} \tag{4}$$

$$Precision = \frac{tp_i}{tp_i + fp_i}(i = 1, 2, 3, \ldots \ldots, c) \tag{5}$$

$$Precision\,(macro\,avg.) = \frac{\sum_{i=1}^{c} \frac{tp_i}{tp_i + fp_i}}{c} \tag{6}$$

$$Recall = \frac{tp_i}{tp_i + fn_i}(i = 1, 2, 3, \ldots \ldots, c) \tag{7}$$

$$Recall\,(macroavg.) = \frac{\sum_{i=1}^{c} \frac{tp_i}{tp_i + fn_i}}{c} \tag{8}$$

$$F1 - score = \frac{2 \times Precision\,(macro\,avg.) \times Recall\,(macro\,avg.)}{Precision\,(macro\,avg.) + Recall\,(macro\,avg.)} \tag{9}$$

In the above, Eq. (5) and (7) calculate *Precision* and *Recall* of an individual class, respectively. Equations (4), (6), (8), and (9) calculate a macro average of the metrics (an average of the metric calculated for all the classes). The macro averaging method treats all the types equally irrespective of the class size (Sokolova and Lapalme 2009).

As we can observe from Eq. (4) to (9), *Accuracy* indicates the average class fraction of 'total correct predictions' among the 'total number of predictions.' This metric provides an excellent indication of the overall performance of a classifier. *Precision* indicates the rate of 'correct' positive class predictions among all the positive predictions made (*tp* ratio among all the predicted *p*). At the same time, *Recall*

Fig. 8 *Accuracy* (%) and *F1-score* (%) for tenfold cross-validation on "train set"
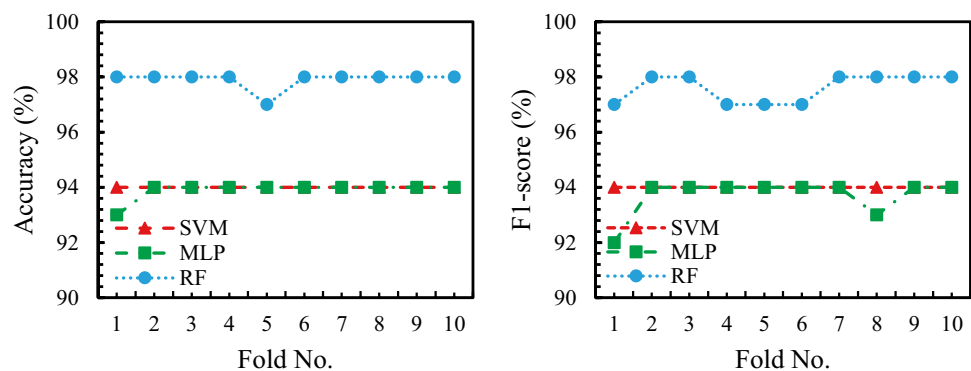
indicates the rate of 'correct' positive predictions among all the positive predictions that could have been made (tp ratio among all the t). *F1-score* is a metric that calculates the mean of *Precision* and *Recall* in a harmonic manner, which can penalize any extreme value of either metric (Géron 2019).
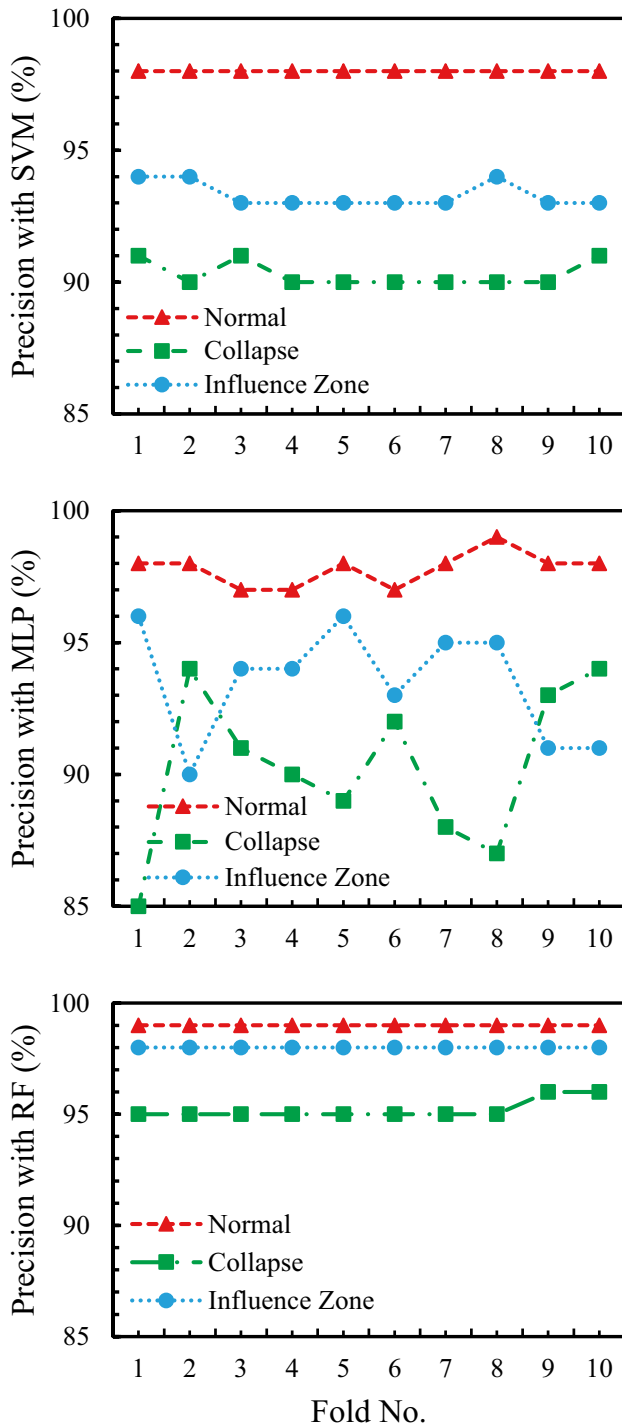

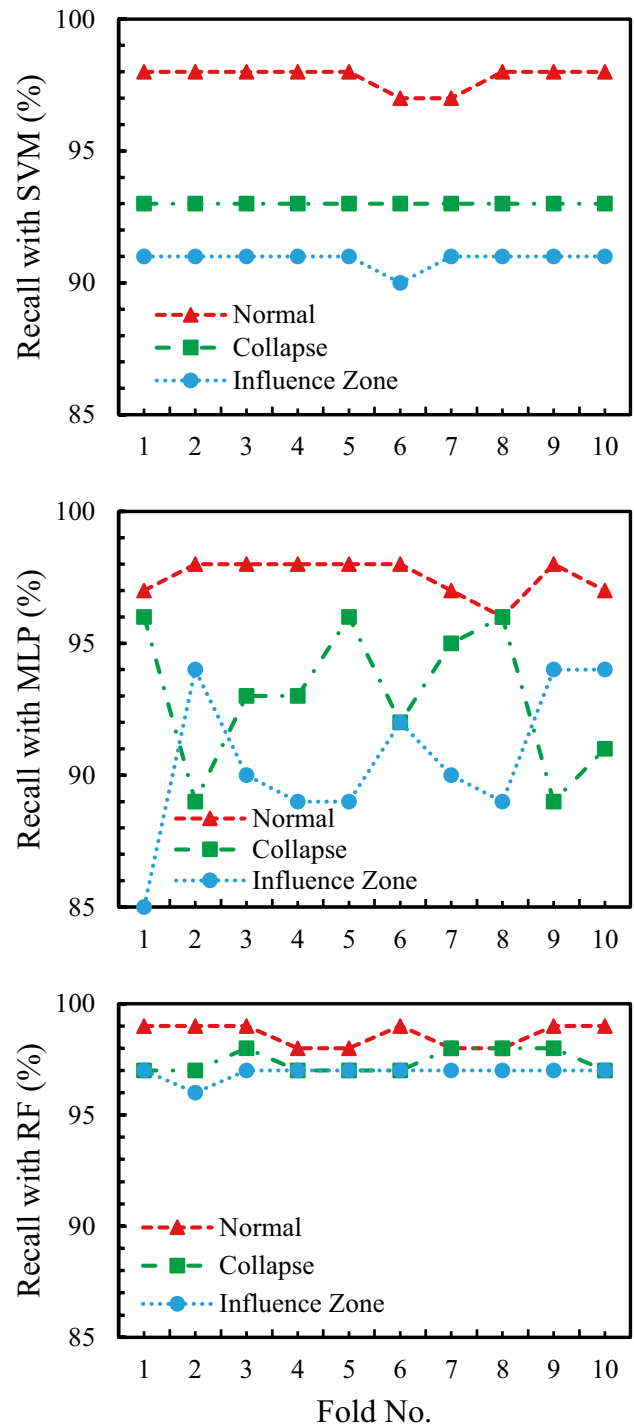**Fig. 9** *Precision* (%) values for tenfold cross validation

Fig. 10 *Recall* (%) values for tenfold cross-validation

6 Results and Analysis

6.1 Classifier Performances

The *Accuracy* and *F1-score* of the three classifiers trained by tenfold cross-validation are displayed in Fig. 8. The

Accuracy measures are almost identical for the different bootstrap samples within a classifier, indicating that the samples are almost similar and representative of the "train set" population. The *Accuracy* range for all the three classifiers is between 93 to 98%, indicating excellent overall performance. The *F1-score*s vary between 94 and 98% between classifiers for ten folds. The maximum *Accuracy* and *F1-score* on the "train set" were achieved by the RF classifier.

The *Precision* and *Recall* measures on the tenfold cross-validation for the three classes are shown in Figs. 9 and 10, respectively. The *Precision* measures for all the classes and classifiers vary between 88 to 99%, indicating overall excellent positive prediction accuracy. It is observed that the 'Normal' class has the maximum *Precision* and *Recall* values for all the three classifiers for individual classes. On the other hand, 'Collapse' has less *Precision* but more *Recall* than 'Influence Zone' for most of the folds of the classifiers. Although it is not desirable to have a lower *Precision* for 'Collapse,' falsely classifying 'Collapse' as 'Influence Zone' is less harmful than classifying that as 'Normal'. It is expected that the 'Normal' class will be the best-predicted class as it has an abundance of data along the tunneling alignment, making it more representative of the population. The *Precision* and *Recall* measures for all the classes are more than 85%, ensuring effectiveness in classifying each class. For the RF classifier, the *Precision* and *Recall* measures for the three classes are at least 95% and 96%, respectively.

The prediction results on the "test set" are shown in Tables 7 and 8. Table 7 shows the results in confusion matrices for the three classifiers. The high diagonal values indicate the excellent performances of the classifiers on the "test set." Table 7 shows that the RF classifier provides an *Accuracy* of 97% on the "test set," which indicates

Table 7 Confusion matrices on "test set"

Classifier	True class	Predicted class			Support
		Influence Zone	Collapse	Normal	
SVM	Influence Zone	22,536	1820	363	24,719
	Collapse	1279	21,715	282	23,276
	Normal	238	486	30,681	31,405
MLP	Influence Zone	23,134	1363	222	24,719
	Collapse	1673	21,237	366	23,276
	Normal	419	469	30,517	31,405
RF	Influence Zone	23,674	930	115	24,719
	Collapse	466	22,586	224	23,276
	Normal	111	361	30,933	31,405

Table 8 Prediction results on "test set"

Classifier	Class	Precision	Recall	F1-score	Accuracy
SVM	Influence Zone	0.94	0.91	0.92	
	Collapse	0.9	0.93	0.92	0.94
	Normal	0.98	0.98	0.98	
MLP	Influence Zone	0.92	0.94	0.93	
	Collapse	0.92	0.91	0.92	0.94
	Normal	0.98	0.97	0.98	
RF	Influence Zone	0.98	0.96	0.97	
	Collapse	0.95	0.97	0.96	0.97
	Normal	0.99	0.98	0.99	

outstanding performance (almost equal to the "train set" *Accuracy*). Overall, no overfitting issue was observed for training and testing the three classifiers (e.g., the difference in *Accuracy* between "train set" and "test set" was minimal). It can be seen that the number of 'Influence Zone'
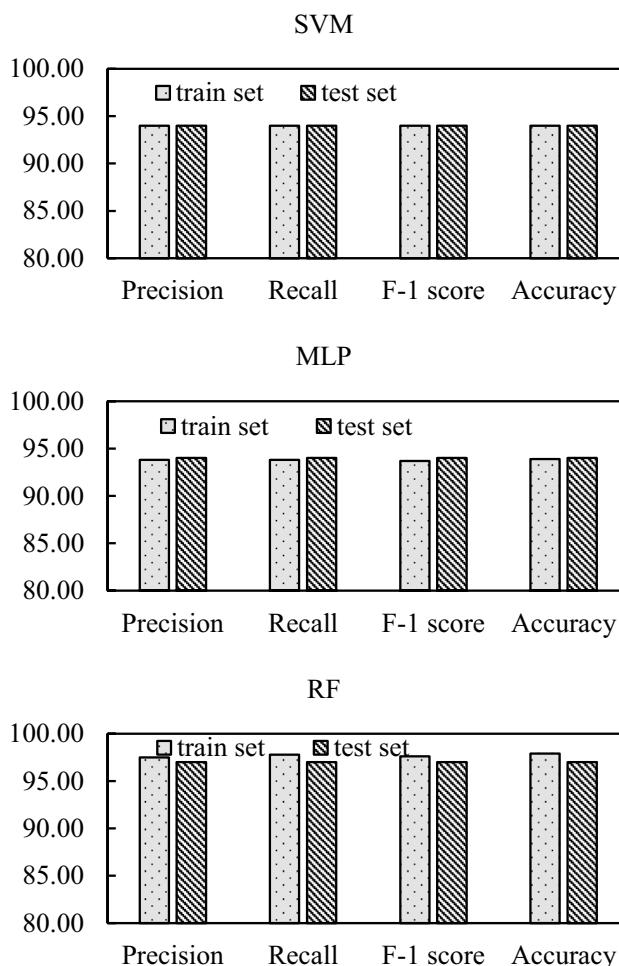

Fig. 11 Machine learning prediction results for "train set" and "test set" data

predicted as 'Collapse' is higher than vice versa in the case of SVM and RF. However, MLP performs differently. Consequently, for 'Influence Zone,' the *Precision* is higher than 'Collapse' for SVM and RF but equal for MLP. Similarly, 'Collapse' has higher *Recall* than 'Influence Zone' for SVM and RF but lower *Recall* for MLP. Overall, RF outperforms the other classifiers in all the performance metrics for the three classes. The comparison of the performance metrics of the classifiers on the "train set" and "test set" are shown in Fig. 11. Hence, based on the overall assessment of the three classifiers, RF is proposed for the prediction model.

While choosing a machine learning classifier, the No-Free-Lunch theorem (Wolpert 1996) states that one cannot prefer a specific classifier over the others without making assumptions, and the only way to know which one will work best is to evaluate them all. However, Fernández-Delgado et al. (2014) provided the results of classification tasks performed over diverse types of data sets by 179 different classifiers. They showed that the classifier families belonging to RF, SVM, and MLP perform better than the others in the majority of the scenarios. They also showed that RF typically performs the best among these three families, followed by SVM for most of the data types. In the current study, RF and MLP classifiers performed well accordingly, which was not the case for SVM.

In this study, the data set utilized is nonlinear, scattered, and prone to noise as the TBM has some vibration accompanying the variability in the ground condition. SVM performs better with a large number of features having fewer data instances. However, in the case of many data instances, the Gaussian *rbf* kernel of SVM, which deals with nonlinear classification problems by adding similarity features, becomes computationally expensive (Géron 2019). In addition, *rbf* is a low-bias kernel, which can be substantially affected by noise inside the data set and perform poorly. On the other hand, RF performs better with a large data set having an excellent tolerance to noises (Zhu et al. 2021). SVM is a strictly binary classifier, whereas RF and MLP are inherently multiclass classifiers, making them more suitable for the current problem study.

Moreover, both RF and MLP can adapt better to highly nonlinear data sets. Structural risk minimization (utilized by SVM) is better than empirical risk minimization (utilized by MLP) to handle overfitting. However, proper tuning of the hyperparameters of MLP (e.g., number of hidden layers, solver, activation function) can alleviate this issue. Also, the vast number of data instances in the current data set is suitable for the MLP classifier to fine-tune the network architecture accordingly. In this study, the hyperparameters tuned through the grid search method (Table 5) successfully alleviated the issue of overfitting. Hence, the above discussions provided insights into why SVM could not perform as expected compared to the other classifiers.

6.2 Multicollinearity

In the case of machine learning, using collinear variables as input features might adversely affect the model performance (Garg and Tai 2013; Li et al. 2021). However, for RF, the effect of multicollinearity is trivial to the prediction performance (Zhan et al. 2018). That being said, having collinear features can cause RF to misjudge the importance of its input features (Sect. 6.3). Also, MLP and SVM have sensitivity towards collinear features for their prediction performance. Hence, to take this issue into account, pairwise correlation coefficients were plotted for the input features utilized in this study, as shown in Fig. 12. The coefficients were plotted for the whole data set as well as data set segments belonging to each of the target feature labels ("Normal," "Collapse," and "Influence Zone"). The intention was to explore if the pairwise correlations differed between the data set segments, which might be informative for the machine learning models.

As we can see, the features E_S and *TPI* exhibit maximum collinearity (correlation coefficient, $r = 1$) in all data set segments. On the other hand, *FPI* has a very high degree of collinearity with both E_S and *TPI* for both the "Collapse" ($r = 0.98$) and "Influence Zone" ($r = 0.97$) data, but not as high in the case of "Normal" data ($r = 0.77$). Finally, there is a reasonable degree of correlation between *F* and *T* (for "Collapse" data, $r = 0.93$; for All data, $r = 0.84$).

The Variable Inflation Factor (*VIF*) was used to explore how these collinearities may have affected the prediction results by the classifiers. In this process, the most inflated factors were iteratively removed to check on their impact on the models. At first, we removed the feature E_S (between the 100% collinear feature E_S and *TPI*) and retrained the models (RF, SVM, MLP). The newly trained models showed no difference in their performance from their original versions. However, further removal of any other potentially collinear features (e.g., *FPI, F*) resulted in a drop of all performance metrics across all the classifiers (e.g., *Accuracy* for RF, MLP, and SVM were reduced by 1.03%, 3.19%, and 4.25%, respectively, on "test set"). This was somewhat expected, as we can observe the loss of information to some degree for removing the features *FPI* and *F* having correlation coefficients, $r < 1$. Also, the high collinearity is not similar across the data set segments for the same pair of elements. Thus, discarding a feature because of high collinearity in the whole data set might result in loss of information existing in individual data set segments, where the collinearity is not so high. Therefore, without a penalty to prediction *Accuracy*, we could only remove one input feature (E_S). Also, the classifiers' performances were not affected before and after the feature removal, which indicated no effect of multicollinearity in the models' prediction performance.
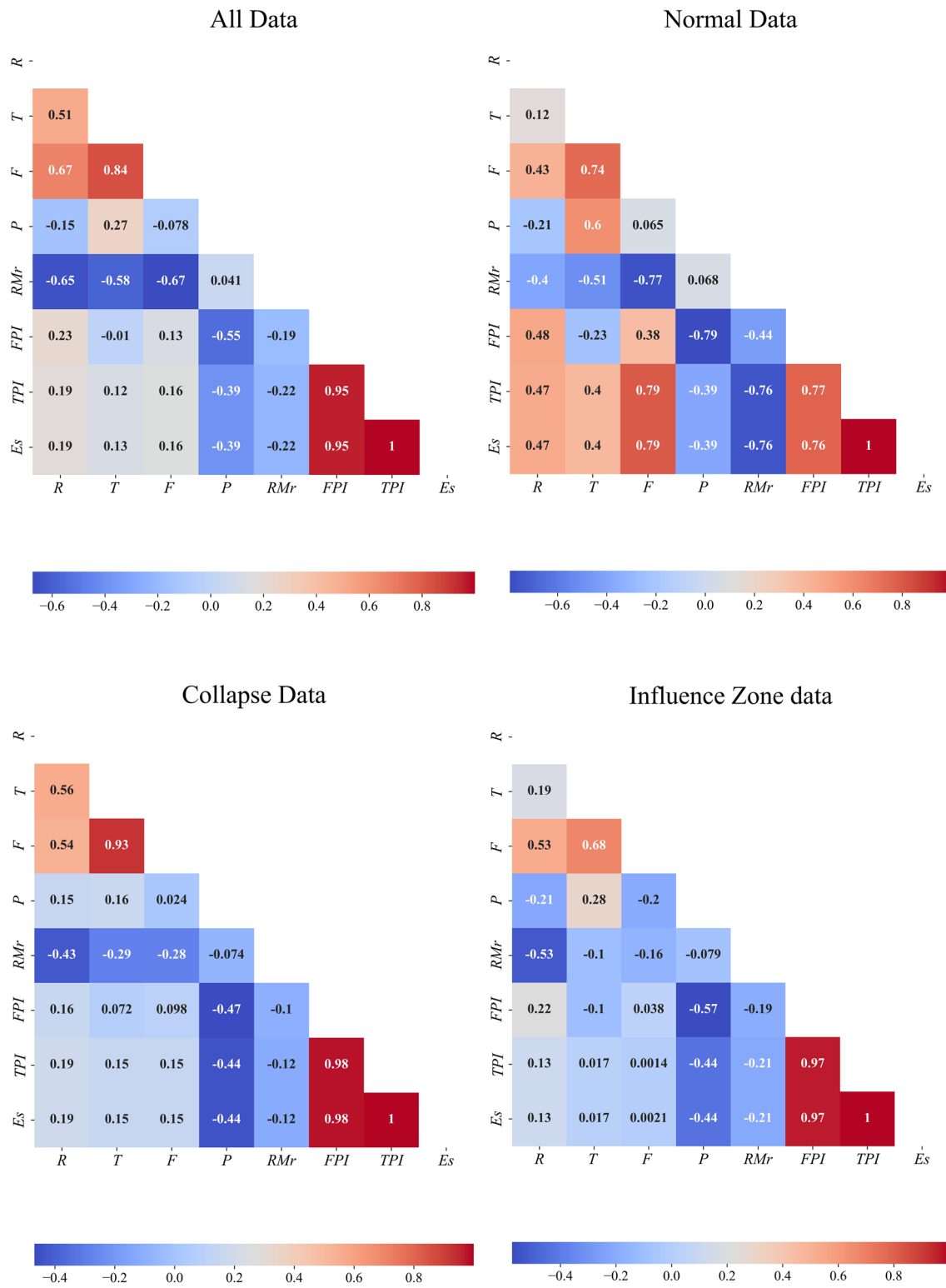
All Data

Normal Data

Collapse Data

Influence Zone data

Fig. 12 Correlation coefficients (*r*) representing pairwise correlation between the input features for, All Data, Normal Data, Collapse Data, and Influence Zone Data

As mentioned earlier, the collinear features might impact the assessment of feature importance by the RF classifier. Hence, the following section calculates feature importance by RF classifier for the input features after removing E_S.

6.3 Feature Importance

As can be concluded, the random forest classifier is the best-performing classifier for the current research. Hence, we analyzed the contribution of each feature variable on the model's performance based on a unique feature of this classifier, which is Gini impurity (Kelleher et al. 2020; Géron 2019).

Random forest is an ensemble learning algorithm built with multiple decision trees. For a classification problem, each decision tree splits the data at each node by utilizing the value of an input feature and creating pure subsets. At each node, a feature variable is selected such that it optimally splits the data into homogenous groups and minimizes the level of impurity at the next node. The measure of how well a node has split the data can be obtained through the Gini index. The Gini index can be understood by how often the target levels of the instances in a data set would be misclassified if predictions were made based only on the distribution of the target levels in the data set (Kelleher et al. 2020). The Gini index of a data set D is measured by

$$Gini(t, D) = 1 - \sum_{l \epsilon levels(t)} (P(t = l))^2 \tag{10}$$

where t is the target or output feature, l is the level of t, and $P(t = l)$ is the probability of target feature $t = l$ for a randomly selected data instance. When the data set D is partitioned using the feature d, a subset of D is created for each level/threshold of that feature. Therefore, the Gini index of the partitioned data set for level k of the feature d is

$$Gini(t, D_{d=k}) = 1 - \sum_{l \epsilon levels(t)} (P_{d=k}(t = l))^2 \tag{11}$$

where $P_{d=k}(t = l)$ is the probability of the target feature $t = l$ for a random data instance inside the subset, where $d = k$. The reduction in Gini impurity by feature d can be obtained by subtracting the weighted sum of the Gini indices of the partitions created by that feature from the initial Gini index. This is known as the Variable Importance Measure (*VIM*) of that feature, which is given by

$$VIM = Gini_{gain}(d, D) = Gini(t, D) - \sum_{k \epsilon levels(d)} Gini(t, D_{d=k}) \frac{|D_{d=k}|}{|D|} \tag{12}$$
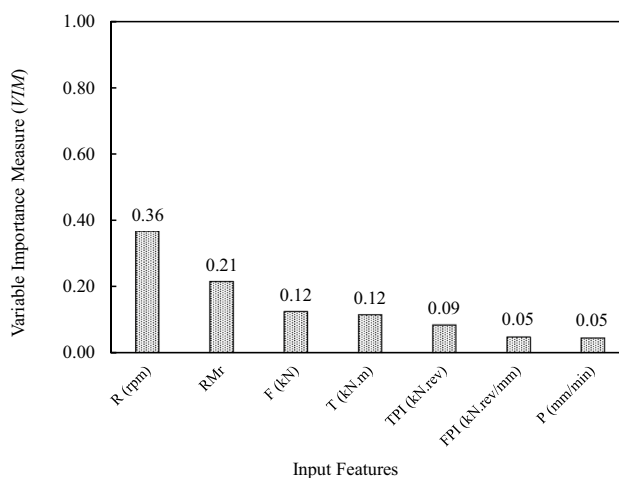

Fig. 13 Variable Importance Measure (*VIM*) of input features by RF classifier

After partitioning the data set, a higher *VIM* indicates a higher reduction of misclassification if trying to predict the target levels of the data set using its distribution alone. In the case of an RF classifier, the total *VIM* of a feature from multiple decision trees are averaged to get the final *VIM* of that feature. Thus, the feature with the highest *VIM* will represent the highest sensitivity of the prediction model towards that feature.

The *VIM* values of the input features calculated by the RF classifier are shown in Fig. 13. The figure shows that the essential feature is R(rpm) having *VIM* of 0.36, whereas the second most important feature is *RMr* (*VIM* = 0.21). F (kN) and T (kN.m) show equal importance with a *VIM* of 0.12, whereas *TPI* shows a bit less importance having *VIM* of 0.09. Overall, all the features have *VIM* of at least 0.05 (5%), indicating satisfactory contribution from all of them to the prediction model. As we can see, three of the TBM operational parameters ($R, T,$ and F) belong to the first four critical features, indicating excellent performance by the direct measurements of the TBM sensors. Another essential feature (2nd best) is the *RMr*. This feature is directly related to the strength and quality of the excavated rock mass. Consequently, it is expected to be a solid indicator of collapse ground condition or its influence zone. Another significant feature could be the lithology of the tunnel alignment. However, there are only two major lithologies (Limestone and Granite) along the tunnel alignment with many collapse incidents, making it a weak feature for the machine learning model (Table 3).

7 Impact of "Influence Zone" Length on Prediction Accuracy

The "Influence Zone" length (*L*/4) utilized in this research approximates the possible zone length around the collapse area. Hence, there is scope to try out different lengths of the proposed "Influence Zone" around each collapse to determine the optimal zone length. Therefore, "Influence Zone" lengths other than *L*/4 were considered to observe the models' performances. The optimal length will be the one that yields the maximum possible zone length with optimal prediction *Accuracy*.

Lengths of *L*/2, 3*L*/4, and *L* in addition to *L*/4 for the "Influence Zone" were considered before and after each "Collapse" zone. Only lengths longer than *L*/4 were
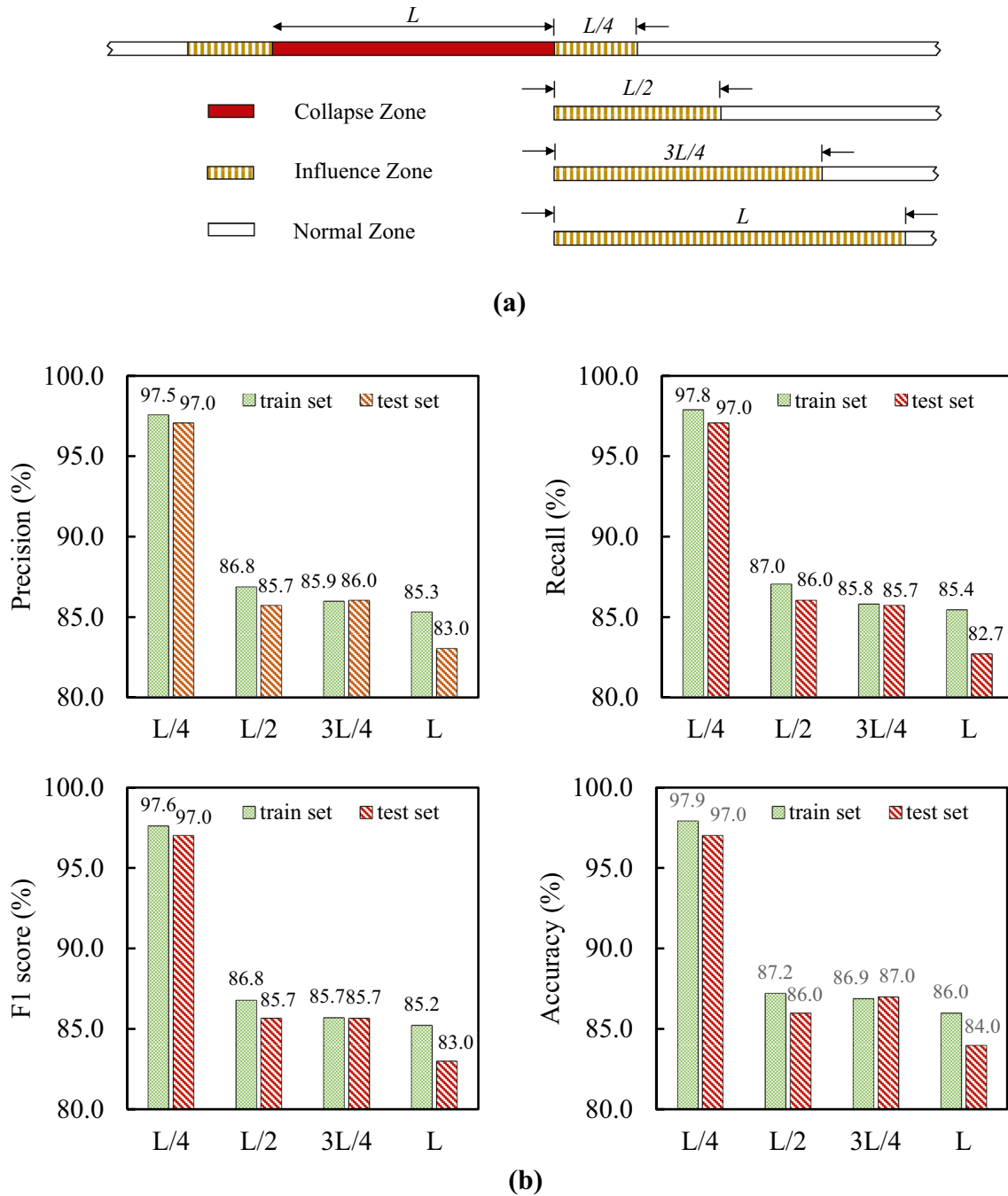
Fig. 14 a Schematic diagram for different "Influence Zone" lengths; **b** Comparison of performance metrics of the prediction models for different "Influence Zone" lengths
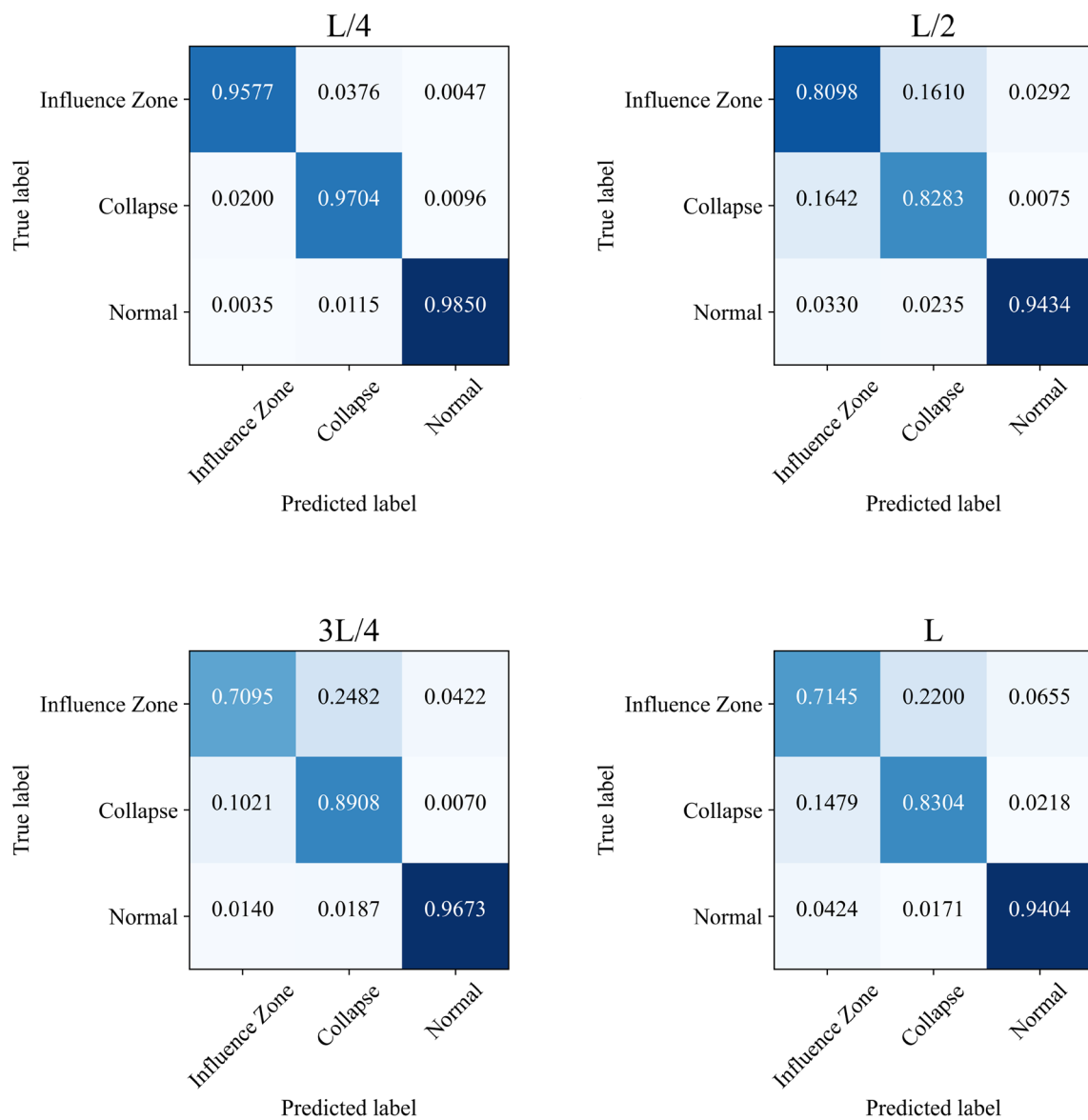
Fig. 15 Normalized confusion matrices for different "Influence Zone" lengths

considered as high prediction *Accuracy* (97%) was already achieved for length *L*/4. It is beneficial to get an increased "Influence Zone" length without sacrificing the prediction performance. Initially, the additional models were trained with similar hyperparameter values (Table 5), resulting in some overfitting issues. Therefore, a new hyperparameter was tuned to handle the problem of overfitting with RF, which is *min samples leaf* (the minimum number of data instances required to be at a *leaf* node). This hyperparameter helps the RF classifier to generalize better and thus avoid overfitting. The value of this hyperparameter was increased from 1 to as high as 4000 to bring the "test set" *Accuracy* nearby to the "train set" *Accuracy*, thus avoiding

the issue of extreme overfitting on the "train set" data. Figure 14a shows the schematic diagram of the four different "Influence Zone" lengths considered, and Fig. 14b shows the performance metrics (*Precision, Recall, F1-score, Accuracy*) on both the "train set" and "test set" data for the four "Influence Zone" lengths considered. In addition, the normalized confusion matrices are also shown in Fig. 15. As we can see, all the performance metrics were over 80%, but they gradually decreased for the increasing lengths of the "Influence Zone" on both the "train set" and "test set" data. Figure 15 shows high diagonal ratios of the normalized confusion matrices for all, with *L*/4 having the highest diagonal ratios among the four.
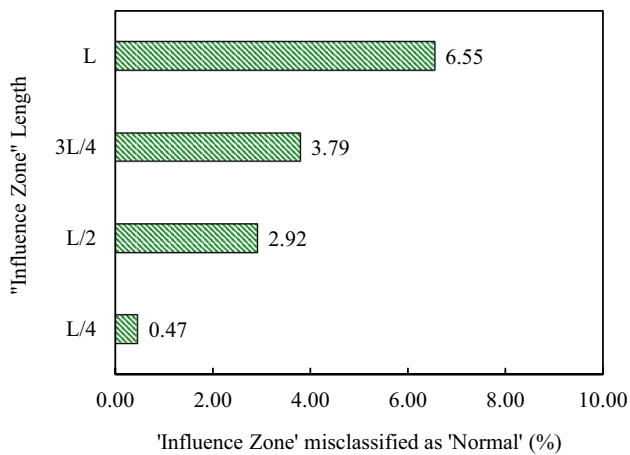
Fig. 16 Percentages (%) of 'Influence Zone' data misclassified as 'Normal' for different "Influence Zone" lengths

The comparison of misclassification of "Influence Zone" as "Normal" for the different lengths is shown in Fig. 16. This figure is important, because the more the assumed length of the "Influence Zone" will be beyond the actual physical range, the more the misclassification of "Influence Zone" instances as "Normal" will be observed. As can be seen, the percentage of misclassification increased from 0.47% to as high as 6.55%, with an increase in the "Influence Zone" length considered.

Deciding an acceptable *Accuracy* to predict the "Influence Zone" during operation is a matter of expert judgment, project demand, and operational conditions. However, from the practical point of view, it is supposed to be the safest for any project that might suffer from frequent collapse incidents to detect the "Influence Zone" as early and accurately as possible with the highest *Accuracy* and the lowest misclassification rate as "Normal." Also, *Accuracy* was reduced by 11% when *L*/4 was increased to *L*/2. A further increase in the "Influence Zone" length resulted in minimal *Accuracy* reduction (by 3%). Hence, it can be inferred that *L*/4 most closely resembled the actual physical range of the "Influence Zone" around collapse, thus posing the least number of threats while performing predictions in the field. Hence, this research suggests this approach for practical use.

8 Application of the Proposed Prediction Model

This prediction model is expected to be utilized by TBM operators performing excavation on rock strata with a high potential to encounter adverse geologic conditions inducive of tunnel boundary collapses. The TBM operator is expected to get predictions of 'Normal' class from the model for most of the tunnel excavation through intact rock conditions without any disturbance. When approaching a "Collapse Zone,"

the prediction model is expected to predict the surrounding ground condition as "Influence Zone" continually for a while. The length of this continual "Influence Zone" prediction will indicate the upcoming "Collapse Zone" length. With a more extended "Influence Zone," the forthcoming "Collapse Zone" is expected to be proportionately longer according to their definition (Fig. 5). Thus, the TBM operator will know if there is any "Collapse Zone" approaching ahead of the existing tunnel face and for how long that specific "Collapse Zone" will continue. When faced with a slightly different operational condition, the current model can serve as a baseline. Then it can train itself further on the newly fed data and potentially adapt itself to make predictions according to the new conditions.

9 Assessment of Features Against Geological Data

A detailed investigation is performed to observe the nature of TBM-related features on different lithologies and *RMr*. The scatter plots of the six TBM-related features of the proposed model concerning the primary two lithologies are shown in Fig. 17. It can be seen that the feature of maximum importance *R* (rpm) is not that different in Limestone and Granite. For *F* (kN), the 'Collapse' and 'Influence Zone' classes mean values are 17% and 42% less in Granite than Limestone, respectively. This difference indicates that their characteristic difference for *F* is more prominent in Granite. For *T* (kN-m), the mean value for 'Collapse' decreases by 52% in Granite compared to Limestone, indicating a characteristic difference for 'Collapse' class in different lithologies. An important thing to notice is the exceptionally high *P* (mm/min) values collected for several data instances of Granite (near the end of both 'Influence Zone' and 'Collapse' class data). These *P* values are in the range of 50–200 mm/min, beyond the normal range of 60–80 mm/min. These data instances are collected from collapse No. 18 (Table 3), representing about 69 m of tunneling operation. The *P* values for this specific 'Collapse' and 'Influence Zone' data were much higher and did not match the other similar data instances. This discrepancy might be why *P* and the related derived features (*FPI, TPI*) were not as important as the other features in the prediction model (Fig. 13). Overall, all the feature values have higher mean values for 'Normal' data instances, except for *P* (mm/min) ('Collapse' and 'Influence Zone' classes have higher *P* in both lithologies). This result is expected as soil condition for both "Collapse Zone" and "Influence Zone" is supposed to be weaker than "Normal Zone." Consequently, penetration should be easier.

It is observed that 'Collapse' only happened in weak rock conditions with *RMr* of IV and V. This explains the

Fig. 17 Scatter plots of TBM operational parameters with lithology (showing data instances of 'Normal', 'Influence Zone', and 'Collapse' classes)
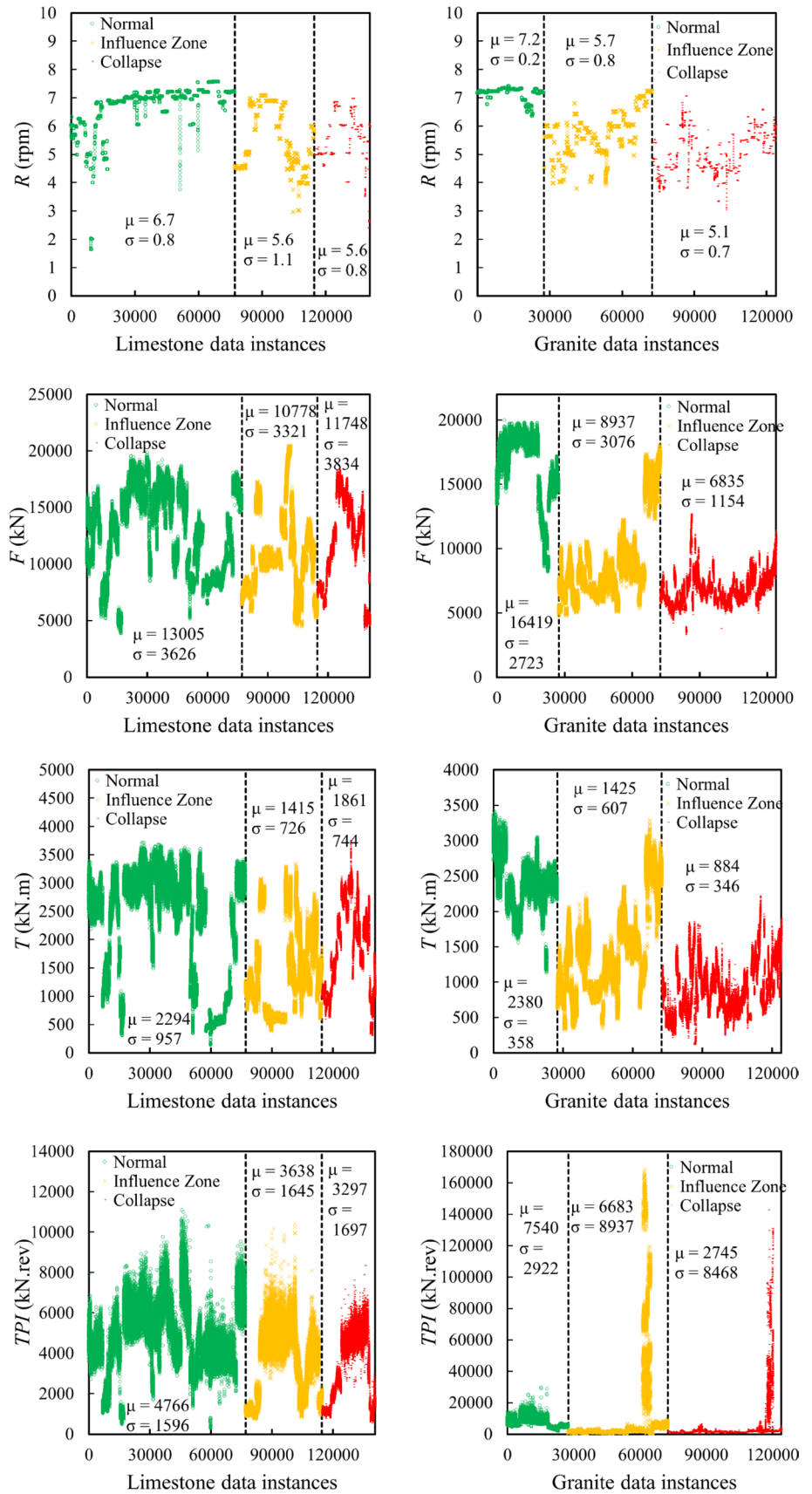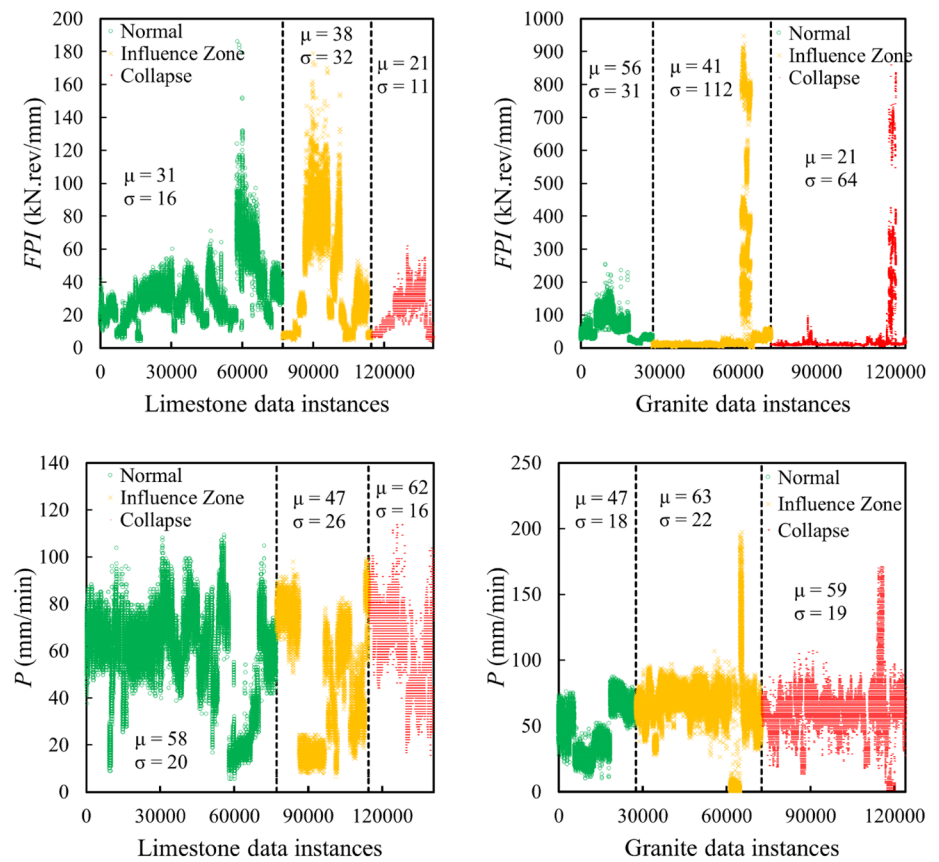
Fig. 17 (continued)

10 Discussion

significance of *RMr* as an input feature (Fig. 13). Hence, it is essential to see how the TBM-related elements behave in different *RMr* ranks. For this reason, histogram plots of these features describing the three classes ('Normal,' 'Collapse,' and 'Influence Zone') against *RMr* Classes III, IV, and V are shown in Fig. 18 (*RMr* Class II has only 'Normal' data instances, hence not presented). It can be noticed that all the three classes ('Normal,' 'Collapse,' and 'Influence Zone') have different peaks and spread under the same *RMr* class in the case of almost all the features. These differences indicate a characteristic variation of the TBM-related features for its class irrespective of the *RMr*. For example, in both *RMr* Class IV and V, the peak of *F* (kN) occurs in a lower value range for 'Collapse' than for both 'Normal' and 'Influence Zone' classes (around 7000 kN in Class IV and 5000 kN in Class V for 'Collapse'). Again, 'Influence Zone' and 'Normal' classes have different peaks and spreads for *F* (kN) in all three *RMr* Classes (higher peak for 'Normal' in *RMr* Class III; for 'Influence Zone' higher peak in *RMr* Class V and multimodal peaks in *RMr* Class IV).

The proposed prediction model utilized both TBM operational parameters, which are direct measurements by the TBM sensors and performance parameters, such as field penetration index (*FPI*) and torque penetration index (*TPI*), which are extracted utilizing additional TBM properties, such as TBM diameter, cutter arrangements, etc. As shown in Sect. 6.3, some of the TBM operational parameters (*R, T,* and *F*) are more effective than the extracted parameters in predicting collapse incidents. Therefore, the model can be considered similarly effective with variation in excavation diameters or TBM properties.

Another important aspect of this study is the efficient use of the TBM data collected during each boring cycle. The data extraction method ensured real-time data collected from each boring cycle. This approach helped capture the actual ground condition as sensed by the TBM sensors. It provided the maximized utilization of the high-frequency data collection practice of TBM, which helped to acquire comprehensive data to represent very few collapse incidents. This, in

Fig. 18 Frequency histograms of TBM-related features against *RMr* classes
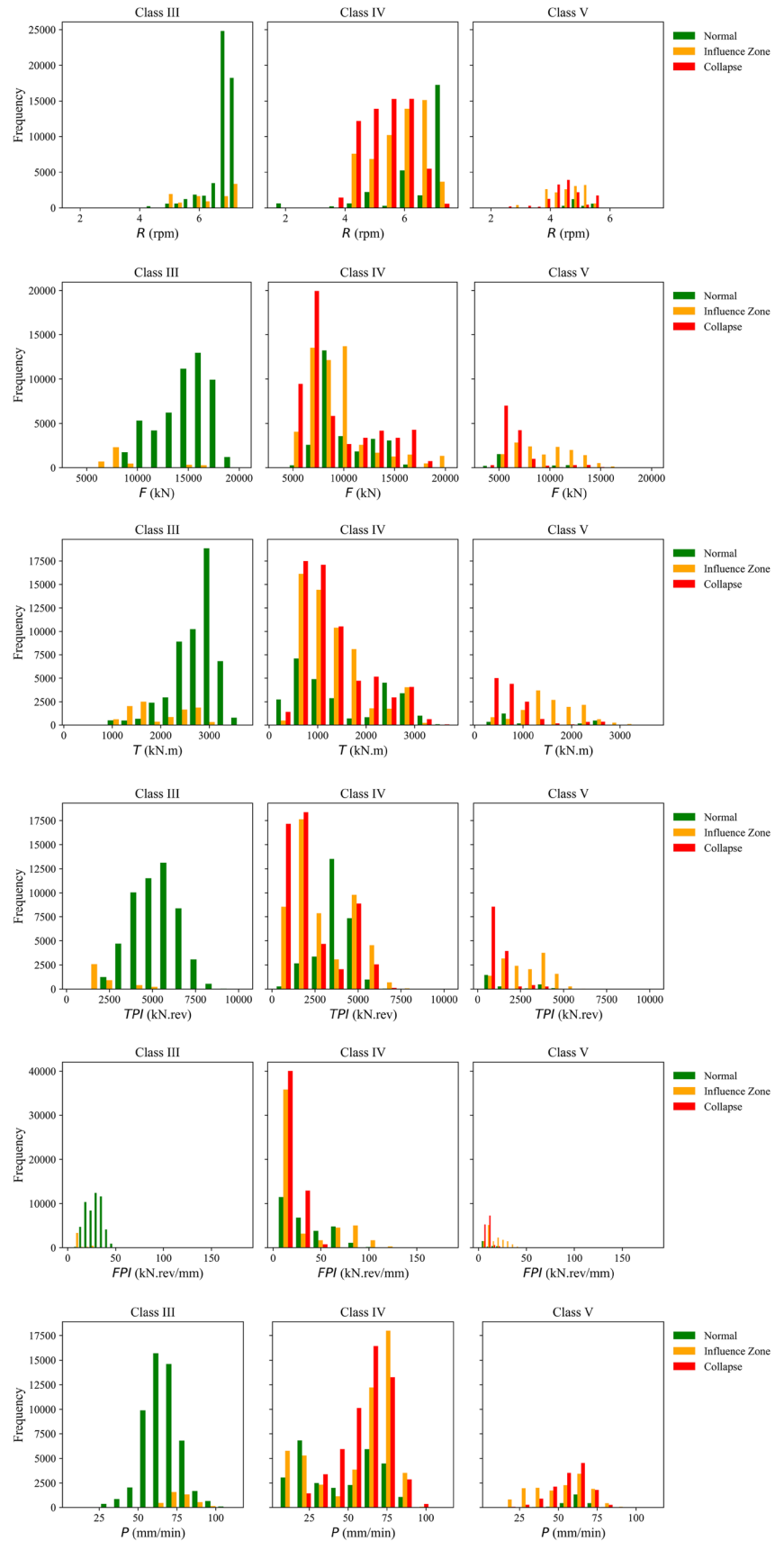
turn, helped the effective training of the machine learning classifiers avoiding data imbalance issues.

The physical understanding of the adverse geologic condition during excavation was efficiently captured by considering the "Influence Zone" in the problem design. Machine learning methods successfully verified the concept by accurately predicting the 'Influence Zone' labels. Overall, this model can be considered beneficial for tunnel engineers experiencing similar ground conditions during operation and researchers trying to simulate a collapsing ground's physical mechanism. The latter can be attributed to the fact that the researchers can gain insights from this study about a new way of understanding and verifying a physical mechanism besides analytical and numerical approaches. The proposed method introduces a more straightforward process of capturing the inherent physical mechanism of collapsing ground without relying on a physical model built on assumptions.

An important aspect to discuss is the use of *RMr* as a geological input feature while predicting collapse ground conditions. It would be ideal to use parameters directly related to the fracture condition of the rock mass (e.g., *RQD*, J_V or J_S) as input features instead of *RMr* to detect collapse ground condition. However, this would require the measurements of these features to be very frequent along the tunnel alignment to have sufficient data points to train machine learning algorithms, which is not practical for a tunneling project. In addition, any individual parameter related to rock fracture does not deliver a suitable generalization of the overall quality of the rock mass (capturing the weakest link). On the other hand, *RMr* provides better abstraction through subjective judgment to the overall quality of the excavated rock mass along the tunnel alignment considering multiple rock mass properties (rock strength, intactness, discontinuity, groundwater condition, etc.). Finally, the utilization of *RMr* as an input feature was justified through the high *Accuracy* achieved by the proposed prediction model and successfully established by *RMr* being the 2nd best feature in detecting the collapse ground condition (Fig. 13). An important aspect about using *RMr* with TBM sensor data as collected in real-time frequency (1 Hz) deserves further attention. Since, *RMr* has its own tolerable range of variation within each assigned class (Liu et al. 2017), it is expected to be independent of the micro-scale variabilities of as-collected TBM sensor data. Hence, the use of on-site TBM data while assigning *RMr* in any form and the use of real-time TBM data for collapse prediction in this study are mutually exclusive.

Alongside the contributions of the current research, some limitations need further improvement. As the model is currently designed, it does not provide any information for the predicted "Influence Zone" instance to know which collapse area it belongs to, which might be ambiguous if two collapses are within proximity (i.e., overlapping "Influence Zone"). Therefore, extra caution should be exercised when the TBM

is passing through the"Influence Zone" belonging to the tail of the current collapse and be prepared to have an upcoming collapse incident until reaching the end of the length expected for the "Influence Zone" belonging to the recent collapse. A multioutput ML model could be trained to understand current conditions to overcome such problems. Another significant limitation is that the proposed model is built to predict collapse incidents happening to specific lithologies and ground. Thus, the model can fail to perform equally well in a significantly different geological condition (e.g., predicting collapse incidents in rock types other than Granite or Limestone). Therefore, data collected on collapses from different projects with various geological conditions could help the model achieve better generalization by removing this bias.

11 Summary and Conclusions

In this research, a prediction model based on machine learning (ML) was proposed to identify the collapse ground conditions ahead of the tunnel face using TBM parameters and rock mass quality in *RMr*. According to the proposed model, the TBM operator can modify TBM control while operating through collapse-prone ground conditions based on the intensity of the expected collapse. Overall, the significant contributions of this research are as follows,

(1) The existence of an "Influence Zone" around the "Collapse Zone" was established and verified by the ML classifier-based prediction model. Identifying these zones ensures the prediction of any upcoming collapse incident with its continuation length during TBM tunneling.

(2) An optimal length for the "Influence Zone" was proposed by assessing the results of the models trained on four different "Influence Zone" lengths considered.

(3) A new data pre-processing technique was proposed for data collected by the TBM data acquisition system. This technique will help generate sufficient data instances to establish ML classifiers with limited data as experienced with collapse ground condition (363 m in total in this case study) against the normal ground condition (more than 19 km in the same case study).

(4) The proposed machine learning prediction model could identify the "Collapse Zone" and "Influence Zone" from the "Normal Zone" ground conditions with at least 97% accuracy based on the "test set" data.

Despite some limitations, the proposed study can provide directions to ongoing tunneling projects experiencing collapse to build ML prediction models with the limited available data. The research also contributed to understanding the collapse mechanism by the prediction success of

the ML-based models and by elaborating the characteristics of the TBM-related features against the tunnel geology (lithology and *RMr*). Overall, the findings of this research are expected to be beneficial to the practicing researchers, tunnel engineers, operators, and experts to interpret better TBM response against collapsing ground conditions and choose appropriate measures of precautions.

Declarations

References

Alimoradi A, Moradzadeh A, Naderi R, Salehi MZ, Etemadi A (2008) Prediction of geological hazardous zones in front of a tunnel face using TSP-203 and artificial neural networks. Tunn Undergr Sp Technol 23:711–717. https://doi.org/10.1016/j.tust.2008.01.001

Benato A, Oreste P (2015) Prediction of penetration per revolution in TBM tunneling as a function of intact rock and rock mass characteristics. Int J Rock Mech Min Sci 74:119–127. https://doi.org/10.1016/j.ijrmms.2014.12.007

Breiman L (2001) Random forests. Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324

Chen J, Li X, Zhu H, Rubin Y (2017) Geostatistical method for inferring RMR ahead of tunnel face excavation using dynamically exposed geological information. Eng Geol 228:214–223. https://doi.org/10.1016/j.enggeo.2017.08.004

Chen Z, Zhang Y, Li J, Li X, Jing L (2021) Diagnosing tunnel collapse sections based on TBM tunneling big data and deep learning: A case study on the Yinsong Project. China Tunn Undergr Sp Technol 108:103700. https://doi.org/10.1016/j.tust.2020.103700

Delisio A, Zhao J (2014) A new model for TBM performance prediction in blocky rock conditions. Tunn Undergr Sp Technol 43:440–452. https://doi.org/10.1016/j.tust.2014.06.004

Ellis K, Kerr J, Godbole S, Lanckriet G, Wing D, Marshall S (2014) A random forest classifier for the prediction of energy expenditure and type of physical activity from wrist and hip accelerometers. Physiol Meas 35:2191–2203. https://doi.org/10.1088/0967-3334/35/11/2191

Exadaktylos G, Stavropoulou M, Xiroudakis G, de Broissia M, Schwarz H (2008) A spatial estimation model for continuous rock mass characterization from the specific energy of a TBM.

Rock Mech Rock Eng 41:797–834. https://doi.org/10.1007/s00603-007-0160-9

Fernández-Delgado M, Cernadas E, Barro S, Amorim D (2014) Do we need hundreds of classifiers to solve real world classification problems? J Mach Learn Res 15(1):3133–3181

Fraldi M, Guarracino F (2009) Limit analysis of collapse mechanisms in cavities and tunnels according to the Hoek-Brown failure criterion. Int J Rock Mech Min Sci 46:665–673. https://doi.org/10.1016/j.ijrmms.2008.09.014

Fraldi M, Guarracino F (2010) Analytical solutions for collapse mechanisms in tunnels with arbitrary cross sections. Int J Solids Struct 47:216–223. https://doi.org/10.1016/j.ijsolstr.2009.09.028

Gao X, Shi M, Song X, Zhang C, Zhang H (2019) Recurrent neural networks for real-time prediction of TBM operating parameters. Autom Constr 98:225–235. https://doi.org/10.1016/j.autcon.2018.11.013

Garg A, Tai K (2013) Comparison of statistical and machine learning methods in modelling of data with multicollinearity. Int J Model Ident Control 18(4):295–312. https://doi.org/10.1504/IJMIC.2013.053535

Géron A (2019) Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media, Inc.

Girmscheid G, Schexnayder C (2003) Tunnel boring machines. Pract Period Struct Des Constr 8:150–163. https://doi.org/10.1061/(asce)1084-0680(2003)8:3(150)

Giussani A (2020) Applied Machine Learning with Python. Egea.

Gong QM, Zhao J (2009) Development of a rock mass characteristics model for TBM penetration rate prediction. Int J Rock Mech Min Sci 46:8–18. https://doi.org/10.1016/j.ijrmms.2008.03.003

Gong Q, Lu J, Xu H, Chen Z, Zhou X, Han B (2021) A modified rock mass classification system for TBM tunnels and tunneling based on the HC method of China. Int J Rock Mech Min Sci 137:104551. https://doi.org/10.1016/j.ijrmms.2020.104551

Guo D, Li J, Jiang S, Li X, and Chen Z (2021) Intelligent assistant driving method for tunnel boring machine based on big data. Acta Geotechnica, 1-12. https://doi.org/10.1007/s11440-021-01327-1

Guo D, Li J, Li X, Li Z, Li P, Chen Z (2022) Advance prediction of collapse for TBM tunneling using deep learning method. Eng Geol 299:106556. https://doi.org/10.1016/j.enggeo.2022.106556

Hassanpour J, Rostami J, Zhao J (2011) A new hard rock TBM performance prediction model for project planning. Tunn Undergr Sp Technol 26:595–603. https://doi.org/10.1016/j.tust.2011.04.004

Hornik K, Stinchcombe M, White H (1989) Multilayer feedforward networks are universal approximators. Neural Netw 2:359–366. https://doi.org/10.1016/0893-6080(89)90020-8

Huang X, Liu Q, Liu H, Zhang P, Pan S, Zhang X, Fang J (2018) Development and in-situ application of a real-time monitoring system for the interaction between TBM and surrounding rock. Tunn Undergr Sp Technol 81:187–208. https://doi.org/10.1016/j.tust.2018.07.018

Huang F, Wang D, Feng Y, and Zhang M (2020) Prediction of the collapse region induced by a concealed karst cave above a deep highway tunnel. Advances in Civil Engineering 2020. https://doi.org/10.1155/2020/8825262

Jeong H, Zhang N, Jeon S (2018) Review of Technical Issues for Shield TBM Tunneling in Difficult Grounds. Tunn Undergr Sp Technol 28:1–24. https://doi.org/10.7474/TUS.2018.28.1.001

Jing L, Li J, Yang C, Chen S, Zhang N, Peng X (2019) A case study of TBM performance prediction using field tunnelling tests in limestone strata. Tunn Undergr Sp Technol 83:364–372. https://doi.org/10.1016/j.tust.2018.10.001

Jung JH, Chung H, Kwon YS, Lee IM (2019) An ANN to Predict Ground Condition ahead of Tunnel Face using TBM Operational

Data. KSCE J Civ Eng 23:3200–3206. https://doi.org/10.1007/s12205-019-1460-9

Kelleher JD, Namee BM, D'Arcy A (2020) Fundamentals of machine learning for predictive data analytics, second edition: algorithms, worked examples, and case studies. MIT Press

Khademi Hamidi J, Shahriar K, Rezai B, Rostami J (2010) Performance prediction of hard rock TBM using Rock Mass Rating (RMR) system. Tunn Undergr Sp Technol 25:333–345. https://doi.org/10.1016/j.tust.2010.01.008

Li S, Liu B, Xu X, Nie L, Liu Z, Song J, Sun H, Chen L, Fan K (2017) An overview of ahead geological prospecting in tunneling. Tunn Undergr Sp Technol 63:69–94. https://doi.org/10.1016/j.tust.2016.12.011

Li S, Nie L, Liu B (2018) The practice of forward prospecting of adverse geology applied to hard rock tbm tunnel construction: the case of the songhua river water conveyance project in the middle of Jilin Province. Engineering 4:131–137. https://doi.org/10.1016/j.eng.2017.12.010

Li J, Li P, Guo D, Li X, Chen Z (2021) Advanced prediction of tunnel boring machine performance based on big data. Geosci Front 12(1):331–338. https://doi.org/10.1016/j.gsf.2020.02.011

Liu Q, Liu J, Pan Y, Kong X, Hong K (2017) A case study of TBM performance prediction using a Chinese rock mass classification system—hydropower classification (HC) method. Tunn Undergr Sp Technol 65:140–154. https://doi.org/10.1016/j.tust.2017.03.002

Liu B, Wang R, Guan Z, Li J, Xu Z, Guo X, Wang Y (2019) Improved support vector regression models for predicting rock mass parameters using tunnel boring machine driving data. Tunn Undergr Sp Technol 91:102958. https://doi.org/10.1016/j.tust.2019.04.014

Liu Q, Wang X, Huang X, Yin X (2020) Prediction model of rock mass class using classification and regression tree integrated AdaBoost algorithm based on TBM driving data. Tunn Undergr Sp Technol 106:103595. https://doi.org/10.1016/j.tust.2020.103595

Macias FJ, Jakobsen PD, Seo Y, Bruland A (2014) Influence of rock mass fracturing on the net penetration rates of hard rock TBMs. Tunn Undergr Sp Technol 44:108–120. https://doi.org/10.1016/j.tust.2014.07.009

Mito Y, Yamamoto T, Shirasagi S, Aoki K (2003) Prediction of the geological condition ahead of the tunnel face in TBM tunnels by geostatistical simulation technique. 10th ISRM Congress 18:833–836.

Mooney MA, Walter B, Frenzel C (2012) Real-time tunnel boring machine monitoring: A state of the art review. North American Tunnelling, 2012 Proceedings, 73–81.

Moreno P, Ruiz M, Gorines FJ (2015) TBM Process Data Management System. Int J Eng Technol 7:431–434. https://doi.org/10.7763/ijet.2015.v7.832

Parise M, De Waele J, Gutierrez F (2008) Engineering and environmental problems in karst—an introduction. Eng Geol 99:91–94. https://doi.org/10.1016/j.enggeo.2007.11.009

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

Rostami J (2016) Performance prediction of hard rock tunnel boring machines (TBMs) in difficult ground. Tunn Undergr Sp Technol 57:173–182. https://doi.org/10.1016/j.tust.2016.01.009

Salimi A, Rostami J, Moormann C, Delisio A (2016) Application of nonlinear regression analysis and artificial intelligence algorithms for performance prediction of hard rock TBMs. Tunn Undergr Sp Technol 58:236–246. https://doi.org/10.1016/j.tust.2016.05.009

Salimi A, Rostami J, Moormann C (2017) Evaluating the suitability of existing rock mass classification systems for TBM performance prediction by using a regression tree. Procedia Eng 191:299–309. https://doi.org/10.1016/j.proeng.2017.05.185

Sapigni M, Berti M, Bethaz E, Busillo A, Cardone G (2002) TBM performance estimation using rock mass classifications. Int J Rock Mech Min Sci 39:771–788. https://doi.org/10.1016/S1365-1609(02)00069-2

Shalev-Shwartz S, Ben-David S (2014) Understanding machine learning: from theory to algorithms. Cambridge University Press. https://doi.org/10.1017/CBO9781107298019

Sheil BB, Suryasentana SK, Mooney MA, Zhu H (2020) Machine learning to inform tunnelling operations: Recent advances and future trends. Proceedings of the Institution of Civil Engineers—Smart Infrastructure and Construction, 1–22. https://doi.org/10.1680/jsmic.20.00011

Sokolova M, Lapalme G (2009) A systematic analysis of performance measures for classification tasks. Inf Process Manage 45:427–437. https://doi.org/10.1016/j.ipm.2009.03.002

Sun W, Shi M, Zhang C, Zhao J, Song X (2018) Dynamic load prediction of tunnel boring machine (TBM) based on heterogeneous in-situ data. Autom Constr 92:23–34. https://doi.org/10.1016/j.autcon.2018.03.030

Wang Y, Jing H, Su H, Xie J (2017) Effect of a fault fracture zone on the stability of tunnel-surrounding rock. Int J Geomech 17:04016135. https://doi.org/10.1061/(ASCE)GM.1943-5622.0000837

Wang Y, Chang H, Wang J, Shi X, Qiu J (2019a) Countermeasures to treat collapse during the construction of road tunnel in fault zone: a case study from the Yezhuping Tunnel in south Qinling, China. Environ Earth Sci 78:1–16. https://doi.org/10.1007/s12665-019-8481-z

Wang Y, Fu N, Fu Z, Lu X, Liao X, Wang H, Qin S (2019b) A semi-automatic coupling geophone for tunnel seismic detection. Sensors 19:3734. https://doi.org/10.3390/s19173734

Wolpert DH (1996) The lack of a priori distinctions between learning algorithms. Neural Comput 8(7):1341–1390. https://doi.org/10.1162/neco.1996.8.7.1341

Wu X, Kumar V, Ross Quinlan J, Ghosh J, Yang Q, Motoda H, McLachlan GJ, Ng A, Liu B, Yu PS, Zhou ZH, Steinbach M, Hand DJ, Steinberg D (2008) Top 10 algorithms in data mining. Knowl Inf Syst 14:1–37. https://doi.org/10.1007/s10115-007-0114-2

Xia Y (2020) Chapter Eleven—correlation and association analyses in microbiome study integrating multiomics in health and disease. In: J Sun (ed) Progress in Molecular Biology and Translational Science. Academic Press, pp 309–491. https://doi.org/10.1016/bs.pmbts.2020.04.003

Xue Y, Li X, Li G, Qiu D, Gong H, and Kong F (2020) An analytical model for assessing soft rock tunnel collapse risk and its engineering application. Geomechanics and Engineering 23(5): 441–454. https://doi.org/10.12989/gae.2020.23.5.441

Yamamoto T, Shirasagi S, Yamamoto S, Mito Y, Aoki K (2003) Evaluation of the geological condition ahead of the tunnel face by geostatistical techniques using TBM driving data. Tunn Undergr Sp Technol 18:213–221. https://doi.org/10.1016/S0886-7798(03)00030-0

Yang X, Li Z, Liu Z, Xiao H (2017) Collapse analysis of tunnel floor in karst area based on Hoek-Brown rock media. J Cent South Univ 24:957–966. https://doi.org/10.1007/s11771-017-3498-5

Yin LJ, Gong QM, Zhao J (2014) Study on rock mass boreability by TBM penetration test under different in situ stress conditions. Tunn Undergr Sp Technol 43:413–425. https://doi.org/10.1016/j.tust.2014.06.002

Zhan Y, Luo Y, Deng X, Zhang K, Zhang M, Grieneisen ML, Di B (2018) Satellite-based estimates of daily NO2 exposure in china using hybrid random forest and spatiotemporal kriging model. Environ Sci Technol 52(7):4180–4189. https://doi.org/10.1021/acs.est.7b05669

Zhang Q, Liu Z, Tan J (2019) Prediction of geological conditions for a tunnel boring machine using big operational data. Autom Constr 100:73–83. https://doi.org/10.1016/j.autcon.2018.12.022

Zhu M, Gutierrez M, Zhu H, Ju JW, Sarna S (2021) Performance evaluation indicator (PEI): a new paradigm to evaluate the competence of machine learning classifiers in predicting rockmass conditions. Adv Eng Inf 47:101232. https://doi.org/10.1016/j.aei.2020.101232

Dong, Guo Jinhui, Li Xu, Li Zhaofeng, Li Pengxi, Li Zuyu, Chen (2022) Advance prediction of collapse for TBM tunneling using deep learning method. Engineering Geology 299106556-10.1016/j.enggeo.2022.106556

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.