

AI-Driven Innovations in Tunnel Construction and Transport: Enhancing Efficiency with Advanced Machine Learning and Robotics



Jagendra Singh¹, Ramendra Singh², Pongkit Ekvitayavetchanukul³, Prabhishek Singh⁴, Manoj Diwakar^{5,6} and Mohd Avesh^{7,*}

¹School of Computer Science Engineering & Technology, Bennett University, Greater Noida, India

²Department of Computer Science and Engineering (IoT), Raj Kumar Goel Institute of Technology, Ghaziabad, India

³The Board of Khonkaen University Affaneirs, Khonkaen University, Thailand

⁴School of Computer Science Engineering and Technology, Bennett University, Greater Noida, India

⁵Department of CSE, Graphic Era Deemed to be University, Dehradun, Uttarakhand, India

⁶Graphic Era Hill University Dehradun, India

⁷Institute of Railway Vehicles and Railway Technology, Technische Universität Dresden, Dresden, Germany

Abstract:

Introduction: Tunnel construction is a high-risk, complex task requiring precision, safety, and efficiency. With growing infrastructure demands, this study proposes a hybrid framework integrating Building Information Modeling (BIM), machine learning models such as Artificial Neural Network (ANN), K-Nearest neighbors (KNN), Support Vector Machines (SVM), and advanced optimization techniques to improve decision-making, predict geological challenges, and automate key operations in large-diameter tunnel projects, enhancing overall project performance and risk management.

Methods: Various methods are employed in the study, including BIM, machine learning, and robust optimization, which can be perceived as enhancing tunnel construction. Prediction using AI-based algorithms, namely ANN, KNN, and SVM, was made possible with real-time sensor data on geological issues. FANUC ROBOGUIDE software was also used to simulate the actions of robots, ensuring that material handling was performed with precision. Among these three, the optimal performance of SVM outshines ANN and KNN.

Results: The results have shown that BIM integrated with machine learning and optimization significantly increased tunnel construction performance. In predicting critical operational parameters, AI-based models, especially SVM, were found to provide an accuracy of 98.56%, outperforming KNN and ANN. Hence, this kind of predictability may allow for real-time modifications in the Tunnel Boring Machines (TBM) settings, thereby decreasing the risks associated with geological uncertainties. Additionally, the FANUC ROBOGUIDE software will ensure more precise and collision-free material handling, further enhancing safety and efficiency in tunnel construction projects.

Discussion: The study demonstrates that integrating BIM with machine learning and robotic simulation significantly enhances tunnel construction efficiency and safety. Among the models evaluated, SVM achieved the highest accuracy (98.56%) in predicting geological challenges. Real-time data processing enabled timely adjustments to TBM operations, while FANUC ROBOGUIDE ensured precise material handling, reducing risks and delays in complex construction environments.

Conclusion: The research currently underway has established the efficacy of integrating BIM, machine learning, and optimization in improving tunnel construction. The applications of AI models, such as SVM, KNN, and ANN, have improved targeted operational parameters and reduced geological risks, with SVM yielding the highest accuracy at 98.56%. Efficiency and safety were further enhanced by real-time data-driven decisions and robotic simulations. The developed framework offers a practical solution for enhancing decision-making and operational efficiency in complex engineering projects.

Keywords: Machine learning, Tunnel construction, Optimization, Robotics simulation.

© 2025 The Author(s). Published by Bentham Open.

This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Public License (CC-BY 4.0), a copy of which is available at: <https://creativecommons.org/licenses/by/4.0/legalcode>. This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

*Address correspondence to this author at the Institute of Railway Vehicles and Railway Technology, Technische Universität Dresden, Dresden, Germany; E-mail: mail2avesh@gmail.com

Cite as: Singh J, Singh R, Ekvitayavetchanukul P, Singh P, Diwakar M, Avesh M. AI-Driven Innovations in Tunnel Construction and Transport: Enhancing Efficiency with Advanced Machine Learning and Robotics. Open Transplant J, 2025; 19: e26671212383139. <http://dx.doi.org/10.2174/0126671212383139250618044241>



Received: January 03, 2025

Revised: March 06, 2025

Accepted: May 02, 2025

Published: July 31, 2025



Send Orders for Reprints to
reprints@benthamscience.net

1. INTRODUCTION

BIM (Building Information Modeling) has gradually become the backbone of the construction industry for managing complex planning in large-diameter tunnel projects. Over the past few decades, BIM has evolved from a simple visualization tool into a comprehensive platform for design coordination, data integration, and project management. BIM now facilitates enhanced collaboration among all stakeholders involved in construction projects. It has become a vital tool in project management by enabling the creation of detailed 3D models that integrate both time (4D) and cost (5D) dimensions [1, 2]. However, despite its widespread adoption in the building sector, BIM remains in the early stages of implementation within tunnel construction, where its full potential has yet to be fully realized.

The literature emphasizes that BIM has the capability to significantly improve the efficiency of tunneling construction work by facilitating better design coordination, enhanced precision in estimating material quantities, and real-time monitoring of worksite progress [3, 4]. It has been observed that the application of BIM decreases errors and rework, resulting in cost savings of 15% or less and the accelerated completion of work within the scheduled timeline. The long-term application of BIM in tunnel construction, however, is challenged by the inherent complexities of underground work, the software's specialized nature and required tools, and the need for integration with emerging technologies such as machine learning and robotics [5, 6].

The application of machine learning (ML) in the building sector represents a significant step towards data-driven decision-making. ML models are increasingly being designed to enhance various building processes, including scheduling, estimation, risk management, and quality control. For tunneling, Artificial Neural Networks (ANN), K-Nearest Neighbors (KNN), and Support Vector Machines (SVM) have been successful in predicting and managing uncertainties in the process of project execution [7, 8, 9]. The models can process large amounts of data, including past and present inputs from sensors, both existing and newly installed, under transient site conditions to predict future disturbances, such as poor geology or the behavior of materials under these conditions.

The key novelty of this research lies in:

- The first-time integration of BIM with AI models and robotics simulation for tunnel construction.
- A data-driven optimization mechanism that updates TBM parameters in real time.
- Use of FANUC ROBOGUIDE software to validate and simulate robotic material handling tasks, ensuring precision and collision avoidance before on-site deployment.

Through this approach, the study bridges the gap between digital modelling, predictive analytics, and physical execution, offering a transformative framework for intelligent and resilient infrastructure development.

2. LITERATURE REVIEW

The literature indicates that Artificial Neural Networks (ANN) have the capability of capturing non-linear relationships in the data and are appropriate for complex building situations where numerous variables behave erratically. K-Nearest Neighbors (KNN) algorithms are straightforward and consistently effective in classification tasks. Support Vector Machines (SVMs), known for their ability to define highly precise decision boundaries, perform best when accurate classification is critical. Machine learning algorithms have been effectively applied in tunnel construction, enabling accurate prediction and real-time adjustment of operational parameters to minimize downtime and avoid costly errors [10, 11].

AI-driven innovations in tunnel construction are presented in the current research paper, highlighting the transformational role of artificial intelligence, machine learning, and robotics in improving process efficiency. Building tunnels is conventionally a complex and time-consuming process. However, AI-driven innovations are improving this landscape, offering better accuracy, speed, and safety in tunneling operations [12]. Most of the critical areas identified in the literature where AI and robotics contributions have been made relate to predictive analytics models in machine learning, which accurately foresee geological conditions for enhanced planning and risk mitigation [13].

With the help of AI algorithms, real-time data provided by sensors integrated into boring machines can be analyzed to optimize cut performance and predict equipment failure, thereby reducing overall downtime. These AI-based systems have made automation in tunnel construction a reality by allowing robotics to take over dangerous and repetitive tasks, thereby improving safety due to the

minimal need for human intervention in hazardous environments [14, 15].

A review of existing studies highlights a growing trend in the adoption of autonomous systems for tunnel construction. Robotics, in particular, is increasingly employed for critical tasks such as drilling, installing support structures, and inspecting completed tunnel segments. The integration of artificial intelligence with robotics enhances system flexibility, allowing for adaptive responses to dynamic and unpredictable construction conditions. However, the literature also points to several challenges, notably the high implementation costs and the need for specialized expertise to operate and maintain these advanced AI and robotic systems [16]. For effective deployment in tunnel construction, an interdisciplinary collaboration among civil engineers, AI experts, and geologists is needed. AI-driven innovations have the potential to enhance the efficiency, safety, and cost-effectiveness of stress processes; however, further research is needed to understand how to overcome the technical and practical challenges associated with their implementation [16, 17].

Robotics is increasingly essential to meeting the demands for precision, efficiency, and safety in modern construction practices. In tunnel projects, in particular, the use of robots is expanding due to their ability to perform repetitive and hazardous tasks with high accuracy. Over time, robotic systems have evolved to support a wide range of applications, from bricklaying to tunnel boring and material handling [18]. When integrated with BIM and machine learning, robotics offers a comprehensive solution to the complexities of tunnel construction. For instance, simulation tools like FANUC ROBOGUIDE enable engineers to test robotic operations under various scenarios, optimize parameters, and anticipate performance outcomes before on-site deployment. This approach not only enhances operational efficiency but also significantly reduces field-related risks [19].

The combination of BIM, machine learning, and robotics is a Smart revolution in tunnel construction. Research has shown an increasing interest in using both technologies together to make construction efficient, facilitate timely decision-making, and make project outcomes more reliable [17]. Several successful examples of these technologies being combined in practice to deliver impressive, faster, safer, and more cost-effective performance are presented through case studies. However, integrating these technologies presents challenges, including the need for specialized expertise, complex data management, and significant initial investment. Nevertheless, the advantages of this combined approach are numerous, and future research is likely to refine and expand its implementation [20].

3. METHODOLOGY

3.1. Methodology and Tools Adopted

The methodology employed here extends established machine learning, Building Information Modeling (BIM), and robotic simulation best practices to address the

challenges of large-diameter tunnel construction. The tools and methods selected are based on recent advances and successful use cases in both civil engineering and smart automation systems.

The application of Artificial Neural Networks (ANN) is justified due to their ability to model non-linear and complex associations between geological factors and TBM performance, making them applicable for predictive modeling in geotechnical engineering. ANN performance is well-established for use in ground settlement prediction and TBM performance analysis; as a result, ANN is an ideal candidate for modeling construction situations under this study.

K-Nearest Neighbors (KNN) was selected for its interpretability and performance for classification based on distance measures. KNN has been successfully used to classify soils and demarcate hazard zones in tunnel construction, based on past geological records. Its capability to identify similar patterns rapidly is critical for sensor-driven decision-making in real-time.

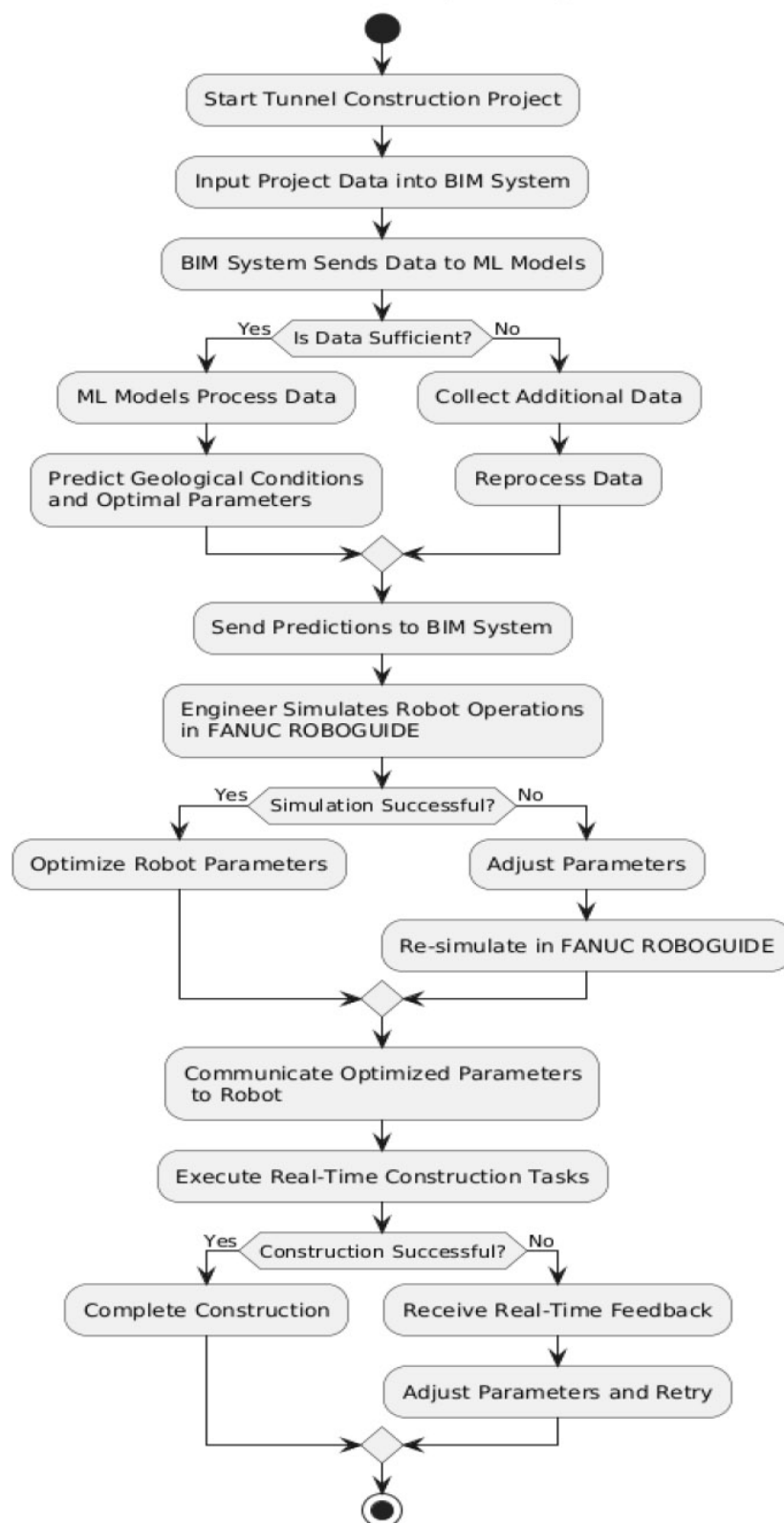
Support Vector Machines (SVMs) were chosen due to their high classification accuracy and stability in high-dimensional spaces. SVM performs well in situations with small but dense datasets, where accurate decision boundaries are pivotal. Previous research has demonstrated the success of SVM in modeling ground conditions and forecasting TBM operating parameters.

Building Information Modeling (BIM) is increasingly recognized as a transformative technology in tunneling and infrastructure development. Its capability to integrate spatial and non-spatial data enables better planning, design coordination, and risk management throughout the construction lifecycle. The integration of BIM in this framework enables centralized data visualization and enhances collaboration among various project stakeholders.

FANUC ROBOGUIDE was utilized to simulate and validate robotic motions. It has gained widespread acceptance in the manufacturing and construction industries for its precision in modeling robot paths, detecting collisions, and offline programming. ROBOGUIDE was found to effectively eliminate on-site errors as well as optimize task performance in tight spaces.

Collectively, these methods and technologies constitute an integrated, research-supported solution to enhance tunnel excavation through predictive analytics, live optimization, and robotic automation.

This study presents a novel interactive machine learning (ML) and robust optimization framework, as illustrated in Fig. (1), for the construction of large-diameter tunnels. The primary objective was to enhance the overall process of constructing the large-scale urban tunnel, which is a challenging task due to geologic variability and dense infrastructure in its vicinity. The study made an effort to predict and counter likely risks, while maximizing operational execution with the assistance of AI-based machine learning algorithms to facilitate informed decision-making throughout the project's lifetime.

Flowchart of the Proposed System**Fig. (1).** Methodology of the proposed system [Source: Authors' own work].

The study utilized multiple machine learning models, including Artificial Neural Networks (ANN), K-Nearest Neighbors (KNN), and Support Vector Machines (SVM). The dataset comprised both historical records and real-time sensor inputs, incorporating factors such as ground conditions, water saturation in subsoil layers, and various geological characteristics relevant to tunnel engineering. These data were used to train the ML models, enabling them to learn and identify underlying patterns. The trained models were then applied to predict potential geological challenges that could impact construction. By processing real-time sensor data, the models were capable of detecting changes in soil composition or the presence of water pockets, factors that could otherwise lead to operational delays or structural issues.

The predictive models identified potential risks, enabling maintenance teams to adjust tunneling operations through an integrated decision-support system. Adjustments included altering the installation sequence, modifying equipment settings, and selecting appropriate materials based on model-driven forecasts. This system's adaptability was crucial for maintaining the construction schedule and ensuring structural stability. The machine learning models were specifically designed to optimize key operational parameters such as speed, pressure, and cutter head configuration of the Tunnel Boring Machines (TBMs). These optimized parameters were then transmitted directly to the robotic platforms, allowing the TBMs to operate in accordance with the predicted geological conditions.

3.2. Machine Learning Models

Artificial Neural Networks (ANNs) are computational models that mimic the processing of neurons in the brain to identify patterns in data through connections between a large number of nodes or neurons. They take input data, apply a transformation using an activation function, and pass the result to the next layer in the network. This structure enables the model to capture complex and meaningful relationships within the data from end to end. At its most basic level, an Artificial Neural Network (ANN) consists of an input layer, one or more hidden layers, and an output layer. Learning involves adjusting the weights of the connections between neurons [17]. This adjustment is carried out through backpropagation, which aims to minimize the prediction error relative to the actual output. It is standard practice to evaluate prediction error using a loss function, such as Mean Squared Error (MSE), as defined in Eq. (1):

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - y^i)^2 \quad (1)$$

Where y_i is the actual value, y^i is the predicted value, and i is the number of samples.

ANN was applied to model and predict the interaction between a range of geological factors with TBM performance in this study. It would predict what a best-case operating profile for head speed, pressure on the cutter face, and cutter head geometry could achieve, based on

first principles and historical data, followed by real-world testing. The network was especially effective at predicting complex geological situations because its ability to capture non-linear relationships provided precision in such circumstances, allowing for adjustments with high agility during tunneling.

K-Nearest Neighbors (KNN) is a simple yet efficient machine learning algorithm used for both classification and regression tasks. K-NN operates on the principle that similar data points will have similar outcomes. The algorithm will use a distance metric, such as Euclidean or Manhattan distance, to find the 'k' closest data points from the training set for every input given. It then takes the votes of neighbors (the majority class in the case of classification/regression), which helps in making a prediction. The Euclidean distance between two points, $x = (x_1, x_2, \dots, x_n)$ and $y = (y_1, y_2, \dots, y_n)$ as defined in Eq. (2):

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2)$$

KNN was used for this research to categorize different geological conditions and foretell the occurrence of risk during tunnel boring. It could quickly retrieve similar past conditions with the same result using historical sensor data, hence making instantaneous predictions. If the geological profile with current sensor readings were similar to a past instance where water pockets caused issues, KNN would predict a high possibility of encountering the same challenges, therefore allowing the construction team to adjust TBM settings proactively. As KNN is interpretable and straightforward, it would be beneficial for making prompt decisions in a complex tunnel construction environment.

SVM (Support Vector Machines) is a type of supervised learning model that can be trained for classification or regression tasks. SVMs work by identifying how to classify points into buckets using a hyperplane that maximizes the margin between classes [4]. In practice, SVMs can utilize a kernel trick to map the data into a high-dimensional space, allowing for the creation of a hyperplane for non-linearly separable datasets. The decision function for SVM is written in Eq. (3):

$$f(x) = \text{sign}(\sum_{i=1}^n \alpha_i y_i K(x_i, x) + b) \quad (3)$$

Where x_i is the support vectors, y_i is the class labels, $K(x_i, x)$ is the kernel function, α_i are the Lagrange multipliers, and b is the bias term.

The SVM method in this study classified the various geological scenarios and predicted optimized TBM operational parameters using real-time data. SVM with a radial basis function (RBF) kernel stands out in its ability to model complex, non-linear relationships between the input features and target variables. The SVM model was justified to accurately predict geological conditions and promote robust optimization of tunnel boring operations, especially in high-dimensional data. SVM has the potential to develop a neat decision boundary, aiding in the accurate prediction of critical operational parameters under various ground conditions.

4. DATASET

The dataset was carefully compiled to comprehensively capture the various factors influencing the construction of large-diameter tunnels within a dense metropolitan environment. It comprised two primary components: historical data and real-time sensor readings. Historical data were sourced from previous tunneling projects in the region, geological surveys, and records of past construction activities. These included detailed information on soil types, proximity to the water table, existing underground utilities, and documented geological events such as landslides and groundwater intrusion. This historical dataset served as a robust foundation for training machine learning models to recognize patterns and predict potential construction-related risks.

In addition to historical data, real-time sensor inputs were collected from Tunnel Boring Machines (TBMs) and various monitoring points along the tunnel route. These sensors captured critical operational parameters, including cutting head pressure, torque, rotational speed, and advance rate, as well as environmental conditions such as soil moisture content, density, and temperature. Continuous real-time data acquisition throughout the tunneling process enabled the machine learning models to dynamically assess the project's current status, facilitating timely updates to predictions and operational recommendations.

4.1. Data Preprocessing

Preprocessing the dataset was a critical step to ensure that the data were clean, accurate, and reflective of real-world conditions in tunnel construction. Given the richness and diversity of the data, ranging from real-time sensor inputs to historical records, a meticulous approach was required to address issues of quality and consistency across heterogeneous sources. The preprocessing workflow included data cleaning, normalization, feature extraction, and dimensionality reduction, along with data augmentation to enhance the robustness of the training sets. These steps were crucial in improving model performance and ensuring that the predictive algorithms produced reliable and highly accurate results.

The initial stage of preprocessing involved thoroughly inspecting the dataset for errors. This included identifying and annotating missing values, correcting anomalies, and removing inconsistencies that could negatively impact model performance. Many of the historical records sourced from geological surveys and previous tunneling projects contained incomplete entries or clerical errors, while some data were found to be skewed. Different imputation strategies were applied based on the nature of the missing data. When a significant portion of a variable was missing, statistical methods such as mean or median imputation were employed, depending on the data distribution. In cases of minor gaps, interpolation techniques were employed to estimate missing values based on adjacent data points.

After defining these points, Z-scores and the Interquartile Range (IQR) were used to identify potential outliers that could influence model predictions. Outliers were then further analyzed to identify those that were more than three standard deviations from the mean or outside of 1.5 times the interquartile range (IQR). These outliers were then checked against the rest of their cluster to confirm whether they were genuine errors or not related to tunneling and were removed. However, even if they represented extreme but valid conditions, the extremes were maintained in order to stress-test our model [12].

Once the data had been cleaned, it was normalized and standardized to ensure that all features contributed fairly to the model training phase. With such a wide variety of sensor data and historical variables, each scaled with different units and found in varied ranges, normalization was essential. The data was further normalized using Min-Max normalization to bring it on a scale of [0, 1]. This is particularly useful when applying machine learning models such as Artificial Neural Networks (ANNs), as the weights often begin with non-zero values. The formula for Min-Max normalization is as shown in Eq. (4):

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (4)$$

Where x is the original value, x_{min} and x_{max} are the minimum and maximum values of the feature, respectively, and x' is the normalized value.

For features exhibiting a Gaussian distribution, standardization was applied to rescale the data to a mean of zero and a standard deviation of one, as shown in Eq. (5):

$$Z = \frac{x - \mu}{\sigma} \quad (5)$$

Where x is the original value, μ is the mean, σ is the standard deviation, and z is the standardized value. Standardization ensured that features with different units or scales did not disproportionately influence the model's learning process.

Feature extraction and feature selection were critical preprocessing steps aimed at reducing dataset complexity while preserving the most informative attributes. The objective was to derive and identify features that could enhance the predictive performance of traditional machine learning models. Domain expertise played a key role in guiding the feature selection process. New features, such as the rate of change in soil moisture and the cumulative pressure exerted on the TBM cutter head over time, were engineered from real-time sensor data. These derived features provided a more accurate representation of geological challenges than the raw attributes alone.

To identify the most relevant features, the study employed feature selection techniques such as Recursive Feature Elimination (RFE) and Principal Component Analysis (PCA). RFE was used to iteratively train models while progressively removing the least important features based on their impact on model accuracy. PCA was applied to reduce the dimensionality of the dataset by trans-

forming the original features into a smaller set of linearly uncorrelated components, while preserving as much of the data's variance as possible. This transformation is represented mathematically in Eq. (6):

$$Z = XW \quad (6)$$

Where X is the original data matrix, W is the matrix of eigenvectors (principal components), and Z is the transformed data matrix.

Data augmentation techniques were then employed to expand the scope of geological conditions beyond those observed in the data, enabling the model to generalize more effectively from rare yet insightful events. Methods like SMOTE (Synthetic Minority Over-sampling Technique) were leveraged to generate synthetic data and balance our dataset, particularly in cases where extreme geological conditions resulted in a rare SOTA, and a data volume 100 times greater was required compared to other scenarios. SMOTE prevents models from developing a bias toward more common conditions by generating synthetic samples, allowing them to be competent across various possible scenarios.

Finally, the cleaned, normalized, and augmented data were consolidated into a single unified dataset for model training. Integrating historical and real-time data required careful alignment to ensure temporal consistency and accurate feature correspondence. For example, sensor readings were aggregated into hourly intervals to match the granularity of the historical data, enabling seamless integration and coherent input representation for the machine learning models.

5. RESULTS AND DISCUSSION

In this study, the primary output consisted of optimized parameters, including speed, pressure, and force, for the tunnel construction process, which served as input settings to guide robotic operations on-site. The FANUC ArcMate 100iD, a versatile serial robot employed for tunnel boring, utilized these parameters to perform its tasks effectively. Prior to physical deployment, the robot's movements were simulated and validated in the FANUC ROBOGUIDE virtual environment, ensuring precision in execution and reducing the risk of operational errors.

The FANUC ArcMate 100iD robot executed tunnel boring operations using parameters optimized through machine learning models. These inputs not only determined the appropriate pressure and force needed for cutting through construction materials but also guided the precise positioning of mechanical components, ensuring smooth operation within the confined and obstacle-laden tunnel environment. By leveraging these ML-derived parameters, the robot was able to dynamically adjust its actions in real time, enabling efficient and safe material handling. The machine learning-powered collision avoidance system significantly minimized the risk of damage to both the robot and construction materials, thereby enhancing overall operational efficiency and safety. The integration of ML-driven optimization with high-fidelity robotic simulation tools, such as FANUC ROBOGUIDE, demonstrated a robust approach for automating critical tasks in large-scale tunnel construction projects.

Once the machine learning models were trained with collected and preprocessed datasets, their predictions about tunnel parameters (optimal) from construction would be further benchmarked in order to evaluate their performance. The accuracy results are shown in Fig. (2).

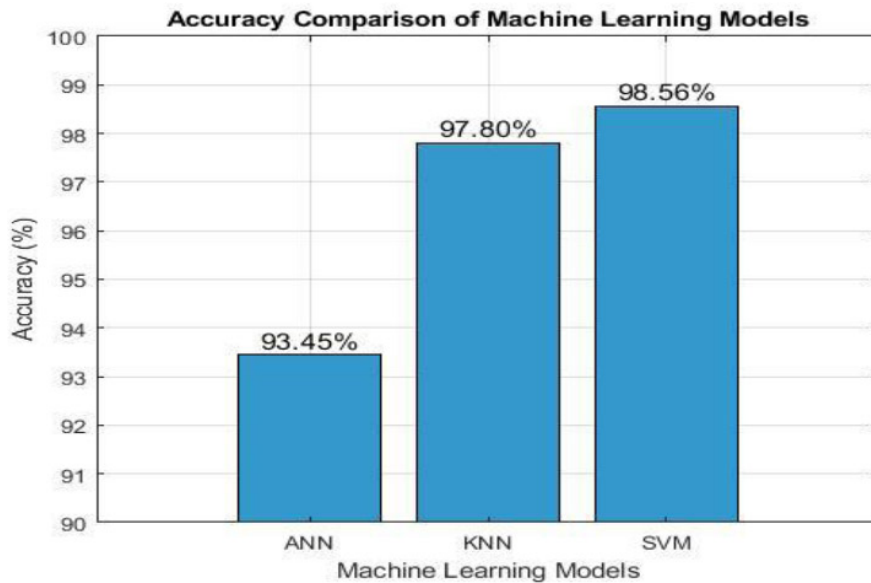


Fig. (2). Accuracy of each used model.

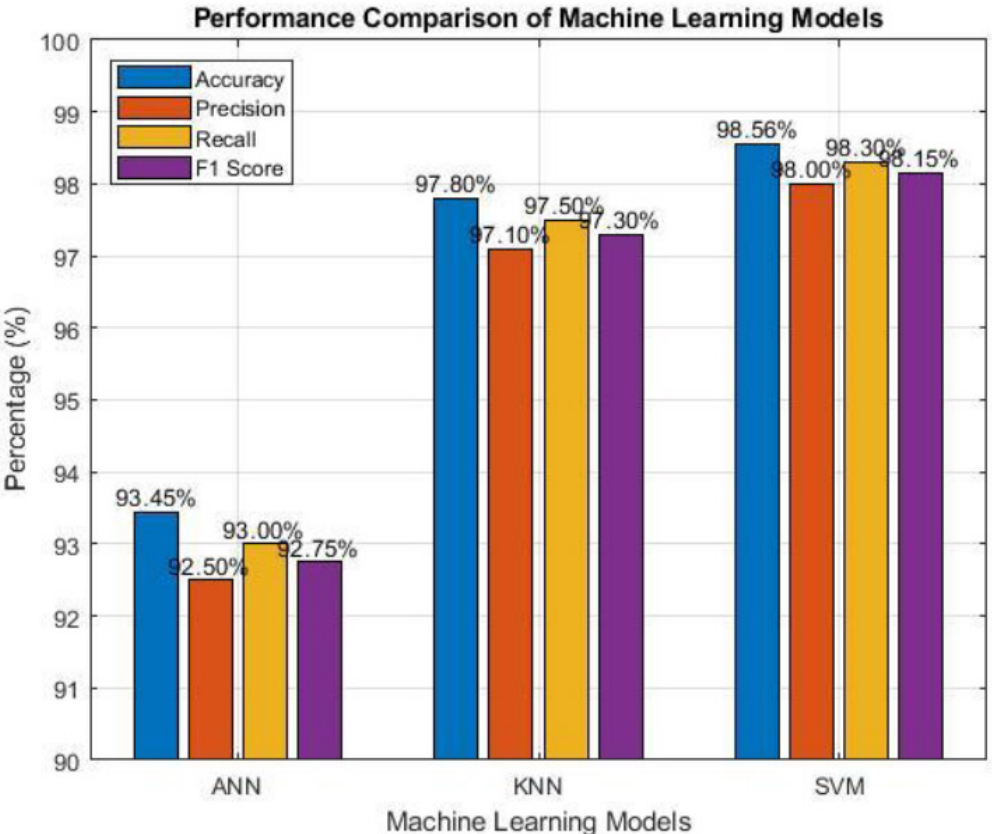


Fig. (3). Performance score of each used model.

The Artificial Neural Network (ANN) model, which flew, scored an accuracy of 93.45%, indicating that the whole idea was indeed a worthwhile endeavor regarding this complex, nonlinear data as well. On the other hand, the K-Nearest Neighbors (KNN) model has shown better performance than the ANN itself, achieving an accuracy score of 97.80%, which is undeniably superior among all machine learning models. It classifies and predicts outcomes based on how closely related they are in the feature space.

Among the three models evaluated, the Support Vector Machine (SVM) achieved the highest performance, with an accuracy of 98.56%. SVM is often preferred for its effectiveness in high-dimensional spaces. The performance metrics of each model are presented in Fig. (3). The Artificial Neural Network (ANN) achieved an accuracy of 93.45%, a precision of 92.50%, a recall of 93.00%, and an F1 score of 92.75%, indicating that it effectively captured complex patterns in the data, albeit with a slight trade-off between precision and recall.

This implies that the model performs well in identifying the actual positive outcome. Another model that performed better than ANN was KNN, with an accuracy of 97.8%, a precision of 97.10%, a recall of 97.50%, and an F1 score of 97.30%. These results reflect the model's ability to accurately match predicted outcomes with actual results.

The Support Vector Machine (SVM) achieved the highest overall performance, with an accuracy of 98.56%, a precision of 98.00%, a recall of 98.30%, and an F1 score of 98.15%. This demonstrates SVM's ability to model the data effectively and form precise decision boundaries.

Fig. (4) presents the confusion matrices for the three machine learning models evaluated: Artificial Neural Network (ANN), K-Nearest Neighbors (KNN), and Support Vector Machine (SVM). For the ANN model, the results show 425 true positives (TP), 25 false positives (FP), 320 true negatives (TN), and 30 false negatives (FN), indicating a balanced but moderate trade-off between precision and recall. In comparison, the KNN model achieved 440 TP, 15 FP, 330 TN, and 20 FN, demonstrating improved classification accuracy with fewer misclassifications relative to the ANN model.

The KNN model demonstrated superior performance by achieving a lower rate of false positives while maintaining a comparable level of false negatives, all while preserving high accurate positive accuracy. However, the SVM model delivered the most balanced results, recording 450 true positives (TP), 10 false positives (FP), 340 true negatives (TN), and only 10 false negatives (FN). This confusion matrix highlights SVM's ability to minimize both false positives and false negatives, making it the most accurate and reliable model among those evaluated.

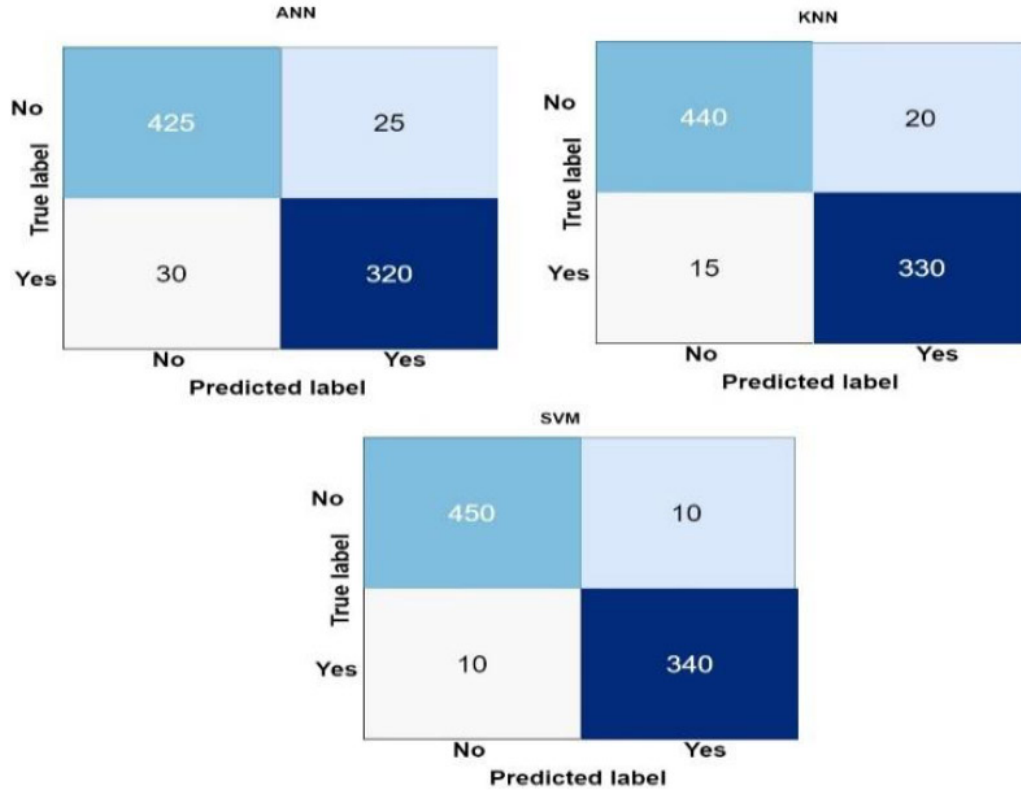


Fig. (4). Confusion matrices of each used model.

Table 1. Optimal parameters for a predefined height from the ML model.

Segment	Place	Optimal Speed (RPM)	Orientation (Degrees)	Alignment (Degrees)	Depth of Installation (meters)
1	A	15	30	10	5
2	B	20	35	12	5.5
3	C	18	32	11	4.8
4	D	22	28	9	5.2
5	E	17	30	10	5.1

Table 1 presents the optimal values for speed, orientation, and alignment of the robot at predefined heights to maintain consistent depth along sinusoidal tunnel paths. These parameters were configured for five tunnel segments to ensure smooth and precise robotic operation in each section, thereby enhancing efficiency and accuracy during the tedious process.

Finally, Fig. (5) illustrates the simulation of the robot executing the placement of the five predefined segments using FANUC ROBOGUIDE software. This simulation played a crucial role in fine-tuning the robot's parameters and optimizing its programming. The process allowed for validation of motion sequences and ensured that the control logic could be effectively transferred to the robot for real-time deployment. As a result, the robot's on-site execution was optimized, contributing to efficient and precise tunnel construction.

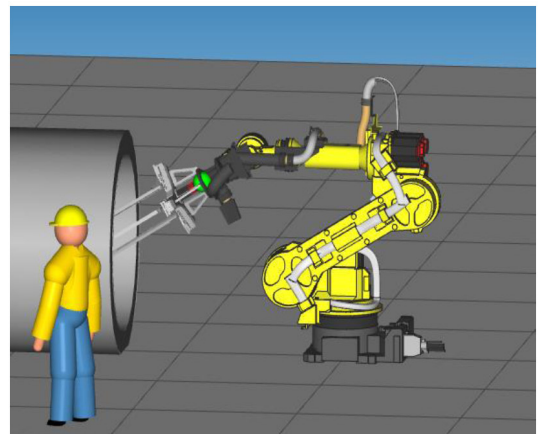


Fig. (5). Robot simulation in FANUC software.

5.1. Sample Size Justification and Statistical Validation

The selection of an appropriate sample size is crucial for ensuring the reliability and statistical significance of AI-driven predictions in tunnel construction. In this study, the sample size was determined based on the availability of historical tunnelling records and real-time sensor data, ensuring that the dataset adequately represented a variety of geological conditions, soil compositions, and underground structural complexities encountered during large-diameter tunnel projects.

5.2. Rationale for Sample Size Selection

The dataset used in this research comprises historical tunneling project data and real-time sensor readings, which include crucial parameters such as TBM operational settings (speed, pressure, and cutter head configuration), geological factors (soil type, moisture content, and water table proximity), and material handling efficiency. The sample size was selected to ensure:

5.2.1. Diversity in Geological Conditions

The dataset encompasses multiple tunnel construction sites, capturing diverse soil and rock formations to enhance model generalization.

5.2.2. Sufficient Data for Machine Learning Model Training

The number of samples was determined to prevent overfitting while ensuring that the AI models (SVM, ANN, KNN) had adequate training data for robust and reliable predictions.

5.2.3. Representation of Real-world Scenarios

The dataset encompasses a range of construction phases, from excavation to material handling, ensuring that the ML models can adapt to various tunneling challenges.

5.3. Statistical Methods for Sample Size Adequacy

To ensure that the selected sample size was sufficient for achieving statistically significant results, the following statistical methods were applied:

5.3.1. Power Analysis

A power analysis was conducted to determine the minimum required sample size for obtaining reliable predictions. The study aimed for a statistical power of 0.80 (80%), ensuring a low probability of Type II errors (false negatives). The analysis considered effect size, significance level ($\alpha = 0.05$), and model complexity to establish an optimal number of data points.

5.3.2. Kaiser-Meyer-Olkin (KMO) Test

The KMO measure of sampling adequacy was used to verify whether the dataset was suitable for training a machine learning model. A KMO value greater than 0.80 indicated that the dataset had sufficient interrelationships among variables, ensuring better feature selection and model training performance.

5.3.3. Bartlett's Test of Sphericity

To further validate dataset adequacy, Bartlett's test of sphericity was performed to assess whether the dataset had significant correlations among features. The test yielded a p -value < 0.05 , confirming that the dataset was suitable for statistical modelling and predictive analysis.

5.3.4. Cross-Validation Strategy

The dataset was divided into training and testing subsets using k-fold cross-validation ($k = 10$) to ensure that the machine learning models generalized well to unseen data. This method reduced the risk of model overfitting and ensured that the predictions remained statistically significant.

The combination of diverse data sources, statistical validation methods, and machine learning evaluation techniques ensures that the sample size selected is adequate for achieving reliable and statistically significant results. Future studies can expand the dataset by incorporating multi-location tunnel construction projects, allowing AI models to generalize across varied geological terrains.

5.4. Advantages and Disadvantages of the Proposed Approach

The integration of Building Information Modelling (BIM), machine learning (ML), and robotics in large-diameter tunnel construction offers several advantages that significantly enhance efficiency, safety, and decision-making. However, like any advanced technological solution, some inherent challenges and limitations must be acknowledged. This section presents a balanced discussion of both the benefits and potential drawbacks of the proposed framework.

5.4.1. Advantages

5.4.1.1. Enhanced Predictive Accuracy for Geological Challenges

- The integration of Artificial Neural Networks (ANN), K-Nearest Neighbors (KNN), and Support Vector Machines (SVM) enables accurate prediction of geological issues such as soil instability and groundwater intrusion.
- The SVM model achieved an accuracy of 98.56%, outperforming both ANN and KNN in classifying geological conditions and optimizing TBM parameters.

5.4.1.2. Real-Time Optimization for Tunnel Boring Machine (TBM) Operations

- The AI-driven framework dynamically adjusts TBM speed, pressure, and cutter head configuration based on real-time sensor data, reducing project delays and mitigating operational risks.
- This predictive capability enables proactive modifications, minimizes unforeseen stoppages, and improves construction efficiency.

5.4.1.3. Improved Safety and Risk Management

- Real-time data analytics help detect anomalies early, reducing the likelihood of tunnel collapses or structural failures.
- Automated decision-making reduces the reliance on manual interventions in hazardous environments, thereby enhancing worker safety.

5.4.1.4. Integration with Robotics for Precision Material Handling

- FANUC ROBOGUIDE software enables collision-free and precise handling of construction materials, improving logistical efficiency.
- Robotic simulations enable engineers to test various tunneling scenarios before deploying them in the real world, thereby reducing errors and material waste.

5.4.1.5. Cost Savings and Sustainability

- Optimizing TBM parameters and implementing predictive maintenance reduces material waste and energy consumption, making the construction process more sustainable.
- AI-driven automation lowers long-term labor costs by reducing the need for manual supervision in repetitive and hazardous tasks.

5.4.2. Disadvantaged

5.4.2.1. High Implementation Costs

- The integration of AI, robotics, and BIM requires substantial initial investment in software, hardware, and specialized training.
- Small-scale projects or companies with limited budgets may face financial constraints in adopting this technology.

5.4.2.2. Dependency on High-Quality Data

- The performance of ML models depends on the availability of extensive, high-quality datasets from geological surveys and real-time sensors.
- Data inconsistencies, missing values, or sensor failures can impact the accuracy and reliability of predictions.

5.4.2.3. Computational Complexity and Real-Time Processing Constraints

- SVM, while achieving the highest accuracy, is computationally intensive, which may lead to increased processing time in real-world applications.
- Deploying real-time AI models in tunnel construction requires robust computational infrastructure, which can be challenging in remote locations.

5.4.3. Interoperability and Integration Challenges

- BIM, ML models, and robotic systems must be seamlessly integrated; however, compatibility issues may arise between different software and hardware platforms.
- Standardized data exchange protocols need to be established to ensure smooth communication between AI-driven TBMs, real-time sensors, and BIM platforms.

5.4.4. Need for Specialized Expertise

- Successful implementation requires expertise in multiple domains, including AI, robotics, civil engineering, and geology.
- A lack of skilled professionals may slow down adoption and increase reliance on external consultants, raising overall project costs.

CONCLUSION

This research introduces a new, integrated approach that combines Building Information Modelling (BIM), machine learning models (ANN, KNN, SVM), and robotic simulation (FANUC ROBOGUIDE) to enhance efficiency, safety, and decision-making during the construction of large-diameter tunnels. By utilizing historical and sensor data during tunnel construction, as well as real-time monitoring, the approach enables the prediction of geological challenges and the dynamic optimization of Tunnel Boring Machine (TBM) operation. Among the models used, the Support Vector Machine (SVM) was found to have a predicted accuracy of 98.56%, outperforming ANN and KNN and being most efficient at finding optimal operation parameters under changing ground conditions.

A significant strength of this research lies in its methodological replicability. The proposed framework, although applied here to tunnel construction, is adaptable to other underground infrastructure projects such as metro rail systems, hydroelectric tunnels, and mining operations, where similar challenges of geological uncertainty, real-time decision-making, and material handling exist. The modular architecture of the framework enables customization based on the type of data available and the operational needs of various engineering domains.

A significant strength of this research is its methodological reproducibility. The framework suggested, while being implemented for tunnels herein, can be generalized to various other underground infrastructure developments, such as metro rail lines, hydro power tunnels, and mining, wherein the same issues of geological uncertainty, dynamic decision-making, and material logistics prevail. The modular design of the framework allows it to be tailored based on available data types and the operational requirements of different engineering domains.

To address these limitations, future research should explore:

- The incorporation of adaptive and lightweight AI models that reduce computational load while maintaining accuracy.
- Sensor fusion and data augmentation techniques to enhance data reliability and expand model training scenarios.
- Reinforcement learning-based control systems for continuous model improvement and autonomous TBM operation adjustments.
- Comprehensive field validation across multiple construction projects and geological contexts to enhance the robustness and scalability of the framework.

In summary, this study presents a transferable and scalable AI-based solution that represents a significant step towards intelligent infrastructure development. By integrating predictive modeling, BIM coordination, and robot simulation, the framework offers novel opportunities for smart, safe, and sustainable tunnel construction, providing a foundation for future developments of AI-enhanced civil engineering systems.

Study Limitations and Future Directions

While this study demonstrates the effectiveness of integrating Building Information Modeling (BIM), Machine Learning (ML), and robotic simulation for optimizing tunnel construction, several limitations should be acknowledged.

Data Dependency and Sensor Accuracy

The proposed framework relies heavily on real-time sensor data for geological predictions and the optimization of TBM (Tunnel Boring Machine) operations. The quality and calibration of sensor inputs directly influence the accuracy and reliability of AI predictions. Any inconsistencies in sensor data, missing values, or environmental interferences may affect the performance of machine learning models. Future research should explore adaptive learning techniques and sensor fusion methods to mitigate data reliability challenges.

Generalization Constraints Across Diverse Geological Conditions

The models used in this study were trained on historical and real-time data specific to the tunnel construction project being studied. While the SVM model achieved high accuracy (98.56%) in predicting geological challenges, its generalization to different geographical locations with distinct soil compositions and underground conditions remains uncertain. Extending the dataset to include diverse geological terrains and conducting field trials in multiple tunnel construction sites would improve the robustness of the proposed approach.

Computational Complexity and Real-Time Execution Challenges

The real-time implementation of AI models, particularly Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs), requires significant computational resources. Integrating ML-driven decision-making

with TBM control systems in real-time may introduce latency issues, especially in high-dimensional data processing. Future studies should explore lightweight AI models, edge computing, or hybrid cloud-based architectures to enhance the system's real-time responsiveness.

Robotic Simulation and Implementation Gaps

While the FANUC ROBOGUIDE simulation validated the proposed robotic operations, translating optimized robotic parameters from simulation to real-world execution remains an area for further investigation. Factors such as hardware limitations, unpredictable environmental changes, and safety considerations during actual tunnel construction require addressing through more extensive field experiments and integration with industry-grade automation platforms.

Economic and Practical Feasibility

The adoption of AI-driven optimization in large-scale tunnel construction necessitates substantial investments in hardware, software, and workforce training. The cost-effectiveness of implementing BIM-integrated ML frameworks needs a comprehensive economic analysis, comparing AI-driven methodologies with conventional construction approaches. Future work should assess the return on investment (ROI) and scalability potential of this technology for widespread industry adoption.

Future Research Directions

To further enhance the applicability of AI-driven innovations in tunnel construction, future research should focus on:

- Developing adaptive and self-learning AI models that continuously improve from real-time data and adjust to evolving geological conditions.
- Expanding datasets to include varied terrains and integrating multi-source data (*e.g.*, seismic activity reports, underground water levels) for better predictive accuracy.
- Exploring reinforcement learning-based optimization for real-time TBM adjustments and automation in tunneling operations.
- Enhancing human-AI collaboration by integrating AI insights with expert decision-making in construction workflows.

By addressing these limitations and advancing research in these areas, the proposed framework can evolve into a more adaptable, reliable, and industry-ready solution for large-diameter tunnel construction.

AUTHORS' CONTRIBUTIONS

The authors confirm their contribution to the paper as follows: M.A.: Methodology; P.E., P.S.: Data collection; J.S., R.S.: Writing the Paper; M.D.: Writing - Reviewing and Editing. All authors reviewed the results and approved the final version of the manuscript.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The data and supportive information is available within the article.

FUNDING

None.

CONFLICT OF INTEREST

The Dr. Mohd Avesh is the Editorial Advisory Board of the journal TOTJ.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

- [1] C. Liu, Z. Wang, H. Liu, J. Cui, X. Huang, L. Ma, and S. Zheng, "Prediction of surface settlement caused by synchronous grouting during shield tunneling in coarse-grained soils: A combined FEM and machine learning approach", *Undergr. Space*, vol. 16, pp. 206-223, 2024.
[<http://dx.doi.org/10.1016/j.undsp.2023.10.001>]
- [2] D. Qiu, X. Li, Y. Xue, K. Fu, W. Zhang, and T. Shao, "Tunnelling and underground space technology incorporating trenchless technology research analysis and prediction of rockburst intensity using improved D-S evidence theory based on multiple machine learning algorithms", *Tunnell. Underg. Space Tech.*, vol. 140, p. 105331, 2023.
[<http://dx.doi.org/10.1016/j.tust.2023.105331>]
- [3] W. Jitchaijaroen, S. Keawsawasvong, W. Wipulanusat, D.R. Kumar, P. Jamsawang, and J. Sunkpho, "Machine learning approaches for stability prediction of rectangular tunnels in natural clays based on MLP and RBF neural networks", *Intell. Sys. Applicat.*, vol. 21, no. January, p. 200329, 2024.
[<http://dx.doi.org/10.1016/j.iswa.2024.200329>]
- [4] X. Li, M. Yao, J. Yuan, Y. Wang, and P. Li, "Deep learning characterization of rock conditions based on tunnel boring machine data", *Undergr. Space*, vol. 12, pp. 89-101, 2023.
[<http://dx.doi.org/10.1016/j.undsp.2022.10.010>]
- [5] H. Huang, J. Chang, D. Zhang, J. Zhang, H. Wu, and G. Li, "Machine learning-based automatic control of tunneling posture of shield machine", *J. Rock Mech. Geotech. Eng.*, vol. 14, no. 4, pp. 1153-1164, 2022.
[<http://dx.doi.org/10.1016/j.jrmge.2022.06.001>]
- [6] X. Yin, F. Gao, J. Wu, X. Huang, Y. Pan, and Q. Liu, "Compressive strength prediction of sprayed concrete lining in tunnel engineering using hybrid machine learning techniques", *Undergr. Space*, vol. 7, no. 5, pp. 928-943, 2022.
[<http://dx.doi.org/10.1016/j.undsp.2022.01.003>]
- [7] X. Tan, W. Chen, T. Zou, J. Yang, and B. Du, "Real-time prediction of mechanical behaviors of underwater shield tunnel structure using machine learning method based on structural health monitoring data", *J. Rock Mech. Geotech. Eng.*, vol. 15, no. 4, pp. 886-895, 2023.
[<http://dx.doi.org/10.1016/j.jrmge.2022.06.015>]
- [8] L. Tang, and S. Na, "Comparison of machine learning methods for ground settlement prediction with different tunneling datasets", *J. Rock Mech. Geotech. Eng.*, vol. 13, no. 6, pp. 1274-1289, 2021.
[<http://dx.doi.org/10.1016/j.jrmge.2021.08.006>]
- [9] M. Liu, S. Liao, Y. Yang, Y. Men, J. He, and Y. Huang, "Tunnel boring machine vibration-based deep learning for the ground identification of working faces", *J. Rock Mech. Geotech. Eng.*, vol. 13, no. 6, pp. 1340-1357, 2021.
[<http://dx.doi.org/10.1016/j.jrmge.2021.09.004>]
- [10] L. Liu, W. Zhou, and M. Gutierrez, "Effectiveness of predicting tunneling-induced ground settlements using machine learning methods with small datasets", *J. Rock Mech. Geotech. Eng.*, vol. 14, no. 4, pp. 1028-1041, 2022.
[<http://dx.doi.org/10.1016/j.jrmge.2021.08.018>]
- [11] H. Park, A. Morisset, M. Kim, H-S. Lee, A. Hessler-Wyser, F-J. Haug, and C. Ballif, "Passivating contact-based tunnel junction Si solar cells using machine learning for tandem cell applications", *Energy and AI*, vol. 14, no. August, p. 100299, 2023.
[<http://dx.doi.org/10.1016/j.egyai.2023.100299>]
- [12] M. Schmitt, "Securing the digital world: Protecting smart infrastructures and digital industries with artificial intelligence (AI)-enabled malware and intrusion detection", *J. Ind. Inf. Integr.*, vol. 36, no. September, p. 100520, 2023.
[<http://dx.doi.org/10.1016/j.jii.2023.100520>]
- [13] G. M, and S.C. Sethuraman, "A comprehensive survey on deep learning based malware detection techniques", *Comput. Sci. Rev.*, vol. 47, p. 100529, 2023.
[<http://dx.doi.org/10.1016/j.cosrev.2022.100529>]
- [14] Z. Ali, S. Saad, K. Rasheed, and S. Ammad, "AI future perspectives and trends in construction", In: *AI in Material Science*, CRC Press: Boca Raton, Florida, 2024, pp. 239-261.
- [15] L. Liu, Z. Song, and X. Li, *Artificial intelligence in tunnel construction: A comprehensive review of hotspots and frontier topics.*, Geohazard Mechanics: Huanggu District, 2023.
- [16] N.R. Kapoor, A. Kumar, A. Kumar, A. Kumar, and H.C. Arora, "Artificial intelligence in civil engineering: An immersive view", In: *Artificial Intelligence Applications for Sustainable Construction.*, Woodhead Publishing: United Kingdom, 2024, pp. 1-74.
[<http://dx.doi.org/10.1016/B978-0-443-13191-2.00009-2>]
- [17] A. Ovid, A. Alsharef, S.M. Jamil Uddin, and A. Albert, "Applied AI and robotics for construction operations—a smart review of the state of the science", *Construct. Res. Cong.*, 2024pp. 913-923
[<http://dx.doi.org/10.1061/9780784485262.093>]
- [18] S.S. Lin, S.L. Shen, A. Zhou, and X.S. Chen, "Smart techniques promoting sustainability in construction engineering and management", *Engineering*, vol. 45, pp. 262-282, 2024.
[<http://dx.doi.org/10.1016/j.eng.2024.08.023>]
- [19] K.C. Onyelowe, F.F. Mojtabehi, A.M. Ebid, A. Rezaei, K.J. Osinubi, A.O. Eberemu, B. Salahudeen, E.W. Gadzama, D. Rezazadeh, H. Jahangir, P. Yohanna, M.E. Onyia, F.E. Jalal, M. Iqbal, C. Ikpa, I.I. Obianyo, and Z.U. Rehman, "Selected AI optimization techniques and applications in geotechnical engineering", *Cogent Eng.*, vol. 10, no. 1, p. 2153419, 2023.
[<http://dx.doi.org/10.1080/23311916.2022.2153419>]
- [20] M.A. Musarat, A.M. Khan, W.S. Alaloul, N. Blas, and S. Ayub, "Automated monitoring innovations for efficient and safe construction practices", *Results Eng.*, vol. 22, p. 102057, 2024.
[<http://dx.doi.org/10.1016/j.rineng.2024.102057>]

DISCLAIMER: The above article has been published, as is, ahead-of-print, to provide early visibility but is not the final version. Major publication processes like copyediting, proofing, typesetting and further review are still to be done and may lead to changes in the final published version, if it is eventually published. All legal disclaimers that apply to the final published article also apply to this ahead-of-print version.