

# Advancing Tunnel Construction Reliability with Automated Artificial Intelligence under Geotechnical and Aleatoric Uncertainties



Jagendra Singh<sup>1</sup>, Prabhishek Singh<sup>1</sup>, Vinayakumar Ravi<sup>6,\*</sup>, Sandeep Kumar<sup>2</sup>, Manoj Diwakar<sup>3,4</sup>, Alanoud Al Mazroa<sup>7</sup> and Indrajeet Gupta<sup>5</sup>

<sup>1</sup>School of Computer Science Engineering & Technology, Bennett University, Greater Noida, India

<sup>2</sup>Department of Computer Science & Engineering, Maharaja Surajmal Institute of Technology, Delhi, India

<sup>3</sup>CSE Department, Graphic Era deemed to be University, Dehradun, Uttarakhand, India

<sup>4</sup>Graphic Era Hill University, Dehradun, Uttarakhand, India

<sup>5</sup>School of Computer Science & AI, SR University, Warangal - 506371, Telangana, India

<sup>6</sup>Center for Artificial Intelligence, Prince Mohammad Bin Fahd University, Khobar, Saudi Arabia

<sup>7</sup>Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University (PNU), P.O. Box 84428, Riyadh11671, Saudi Arabia

## Abstract:

**Aims:** This research seeks to improve the reliability and sustainability of tunnel construction by employing automated AI techniques to manage geotechnical and aleatoric uncertainties. It utilizes machine learning models, including Gradient Boosting Machines (GBM), AdaBoost, Hidden Markov Models (HMM), and Deep Q-Networks for Reinforcement Learning, to predict and reduce environmental impacts. The effectiveness of these algorithms is assessed using various performance metrics to demonstrate their impact on enhancing tunnel construction processes.

**Background:** While tunnel construction is vital for modern infrastructure development, it poses significant environmental challenges. Traditional methods for assessing environmental impacts often rely on manual techniques and overly simplistic models that fail to consider the complex interactions and inherent uncertainties of geotechnical and aleatoric factors. This research aims to overcome these limitations by applying automated AI techniques, particularly machine learning algorithms, to more accurately predict and mitigate environmental impacts.

**Objective:** The goal of this study is to increase the reliability and sustainability of tunnel construction by using AI-based methods to address both aleatoric and geotechnical uncertainties. It focuses on deploying machine learning algorithms such as GBM, AdaBoost, HMM, and Deep Q-Networks for Reinforcement Learning to forecast and manage negative environmental impacts. The algorithms' performance is measured against various criteria to demonstrate their effectiveness in optimizing construction outcomes.

**Methods:** The research applies machine learning techniques, including GBM, AdaBoost, HMM, and Deep Q-Networks, to enhance tunnel construction's reliability and environmental sustainability. These models are designed to predict and mitigate environmental impacts while accounting for geotechnical and aleatoric uncertainties. The models' effectiveness is evaluated using metrics like accuracy, precision, recall, F1 score, log loss, mean squared error (MSE), log-likelihood, cumulative reward, convergence rate, and policy stability, indicating substantial improvements in construction practices.

**Results:** The study shows that using machine learning algorithms significantly enhances tunnel construction reliability and sustainability. GBM achieved a high accuracy of 0.92 and an F1 score of 0.90. Additionally, Deep Q-Networks for Reinforcement Learning effectively identified optimal construction strategies, resulting in a cumulative reward of 950. These outcomes highlight the capability of AI methods to address uncertainties, leading to safer, more resilient infrastructure development.

**Conclusion:** The findings of this research suggest that integrating machine learning algorithms, such as GBM, AdaBoost, HMM, and Deep Q-Networks, substantially improves the reliability and sustainability of tunnel construction projects. These AI approaches effectively manage geotechnical and aleatoric uncertainties, with GBM providing high predictive accuracy and F1 scores and Deep Q-Networks optimizing construction strategies. Adopting these technologies could result in safer, more sustainable, and resilient infrastructure, underscoring their potential for transforming tunnel construction practices.

**Keywords:** Tunnel construction, Environmental impact assessment, Machine learning, Artificial intelligence, Sustainability, Gradient boosting machines, Hidden markov models.

© 2024 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.



Received: July 11, 2024  
Revised: September 01, 2024  
Accepted: September 04, 2024  
Published: October 04, 2024



Send Orders for Reprints to  
[reprints@benthamscience.net](mailto:reprints@benthamscience.net)

\*Address correspondence to this author at the Center for Artificial Intelligence, Prince Mohammad Bin Fahd Univer-sity, Khobar, Saudi Arabia; E-mail: [vinayakumarr77@gmail.com](mailto:vinayakumarr77@gmail.com)

Cite as: Singh J, Singh P, Ravi V, Kumar S, Diwakar M, Al Mazroa A, Gupta I. Advancing Tunnel Construction Reliability with Automated Artificial Intelligence under Geotechnical and Aleatoric Uncertainties. Open Civ Eng J, 2024; 18: e18741495343684. <http://dx.doi.org/10.2174/0118741495343684240910115546>

## 1. INTRODUCTION

Tunnel construction is one of the main components of the development of modern infrastructure, which allows efficient development of transport networks, construction of water supply systems, and commuting underground utilities. However, it frequently results in considerable adverse natural impacts, including disruption of habitats, water pollution with toxic chemicals used in the process of construction, and air pollution. The prominent damages caused to the environment by construction tunnels often lead to significant consequences for ecosystems, biodiversity and humans, which predetermines the necessity of effective mitigations [1-3].

Considering and decreasing these environmental risks is necessary for ensuring sustainable development and lessening negative effects on the environment. The approaches that were traditionally used to assess environmental impacts in tunneling have been based on manual methods and rather simplistic models. Such methods may not demonstrate the involved geotechnical factors and uncertainties typical of aleatoric factors. Therefore, it is possible to state that the authors' conclusion that advanced technologies, such as Artificial Intelligence and Machine Learning, can revolutionize the approaches to how environmental impacts are assessed and designed in tunneling projects is logical and relevant [4-6].

The aim of the provided research is advancing tunnel construction reliability using automated AI techniques under geotechnical and aleatoric uncertainties. The research will focus on the employment of ML algorithms, such as GBM, AdaBoost, HMM, and Deep Q-Networks Reinforcement Learning, to develop predictive models capable of timely assessment and mitigation of environmental risks associated with tunnel construction. The models will be trained and tested with extensive historic tunnel environment data provided by the researchers, which will undergo comprehensive preprocessing, feature detection, and training to maximize model performance.

This research aims to discuss the wider implications of the employment of advanced AI methods as part of the environmental impact assessment EIA processes. Since the research is focused on the use of multiple ML algorithms and corresponding performance evaluation metrics capable of predicting and assessing environmental

impacts, its results and findings may be useful for helping those involved in the construction of tunnels to develop more sustainable and reliable methods of construction that would help mitigate and manage existing and future environmental challenges.

## 2. LITERATURE REVIEW

The application of artificial intelligence and machine learning to numerous aspects of tunnel construction has been a focus of increasing interest. Such interest may stem from the motivations to enhance the reliability, efficiency, and sustainability of tunneling projects in light of the complexities of many geotechnical and environmental conditions. To this end, a number of studies have looked into the applications of ML algorithms to predict TBM performance, predict surface settlement, as well as time series forecasting. A wide variety of AI and ML techniques has been used to model the complex relationships between tunneling characteristics and environmental impacts, including Gradient Boosting Machines, AdaBoost, Hidden Markov Models, and Deep Q-Networks Reinforcement Learning [7-9].

Currently, existing approaches contain several intrinsic drawbacks. First of all, many existing studies employ limited datasets or over-simplified models, which fail to develop adequate generalization and may result in overfitting. Secondly, research in the field requires complete re-evaluation by considering a broader range of metrics for assessing the performances of ML algorithms. The latter may include but are not limited to parameters like accuracy, precision, recall, F1 score, log loss, and mean squared error, which have not been individually discussed for each performance yet. Finally, it is also essential to outline the lack of research dedicated to the implementation of AI and ML in terms of the construction of the entire tunnel [10, 11].

The proposed research addresses some gaps in the literature which has not yet provided a comprehensive framework to apply automated AI to tunnel construction under geotechnical and aleatoric uncertainties. The large dataset of historic tunnel environmental data and the wide range of ML algorithms used in this research to train the models/system are believed to result in a much more robust and more reliable approach to predicting and mitigating environmental impacts. The research will also employ advanced AI techniques like Deep Q-Networks Reinforcement Learning to derive construction strategies

that are optimal based on the environmental conditions [12-16].

The research will be beneficial in terms of devising a number of fresh solutions aimed at enhancing the sustainability and resilience of tunnel construction practices. Finally, through the incorporation of the most innovative AI tools in the process of developing environmental impact assessments, the research can be seen as a groundbreaking one, as it will shed light on new ways of approaching tunnel construction projects as such, from their planning stages to their maintenance [17-19].

### 3. METHODOLOGY

#### 3.1. Data Collection and Preprocessing

Historic tunnel environmental data collected for this research includes a variety of parameters that are of paramount importance for the assessment and mitigation of construction’s environmental impacts. As presented in Table 1, as far as the composition and structure of soil are concerned, the data includes information about the types of soil that can be found along the tunnel’s construction and their key properties that regard texture, porosity, and permeability. This information is necessary to assess the potential risk of soil erosion, sedimentation, and pollution. The data on the quality of rock refers to its strength, fracturing degree and pattern, and weathering state. This information is important to evaluate the stability of the tunnel and assess the risk of rock falls, avalanche formation, and other events that can damage the environment located at the site of construction.

The data that relates most to this problem is the groundwater and seismic activity data. The groundwater data, which includes both water flow and potability, needs to be considered because the company needs to know the physical presence of the groundwater, the process of the water flow in plain view, and flow zones and contaminants in the water to better understand the risk level should the data along with the data on the nearby water bodies. The

seismic activity data contains information about the times and magnitudes of earthquakes in the areas. This data affects the decision to dig the tunnel because the high seismic activity raises the likeliness of an environmental catastrophe in case the tunnel accidentally collapses or malfunctions.

The data collected from the Public Utility Data Repository (PUDR) reveals the quality of both surface water and groundwater. This includes the pH value, turbidity, and the levels of contaminants in the water. This is important baseline data that can be used to understand conditions on site before tunnel construction takes place. Moreover, this type of data is also essential in understanding all potential impacts of the construction project on the local water systems. Additionally, the air quality data shows the concentration of major pollutants of concern, such as PM10, NOx, SOx, and CO2, in the region surrounding the tunnel.

Similar to the water quality data, each type of air quality data is typically collected for the same areas - around the site where the construction will be implemented. A significant addition to the database would be noise and vibration data. This type of data provides detailed reports about baseline levels and how much it is predicted to rise due to construction works. It is important to have this in order to understand the implications for the human population and the wildlife. Data about the atmospheric conditions, the humidity, the ambient pressure, and the temperature that was mentioned above is also important, as it is necessary to know the current environmental condition of the area in which the tunnel is built and the implications of climate change on the ability to build and operate the infrastructure. Finally, data on soil contamination allows the users to know the levels of heavy metals, hydrocarbons in the soil, and other contaminants in order to understand where the risks of soil and water pollution from the construction are greatest.

**Table 1. Dataset information.**

Data Collection Process Parameters	Sample Reading
Soil Composition and Structure	Sandy loam with 60% sand, 30% silt, and 10% clay
Rock Quality	Moderately strong sandstone with some fractures and weathering
Groundwater Levels and Flow	Groundwater table at 10 meters depth, flow rate of 0.5 m/s
Seismic Activity	Magnitude 4.2 earthquake recorded last year
Surface Water Quality	pH 7.2, turbidity 10 NTU, no detectable contaminants
Groundwater Quality	Calcium-magnesium bicarbonate type water, no contaminants detected
Air Quality	PM10: 25 µg/m <sup>3</sup> , NOx: 15 ppb, SOx: 5 ppb, CO2: 400 ppm
Noise Levels	55 dB(A) during the day, 45 dB(A) at night
Vibration Levels	0.2 mm/s peak particle velocity, expected to increase by 50%
Temperature and Humidity	Average temperature: 20°C, average humidity: 60%
Precipitation Patterns	Average annual rainfall: 1000 mm, with most precipitation occurring during the monsoon season
Wind Patterns	Predominantly westerly winds at an average speed of 5 m/s
Soil Contamination Levels	Lead: 50 mg/kg, petroleum hydrocarbons: 100 mg/kg
Sedimentation Rates	Average sedimentation rate: 2 mm/year, expected to increase by 20%
Chemical Spills and Leaks	One minor spill of diesel fuel reported in the last 5 years

### 3.2. Data Cleaning, Normalization, and Feature Engineering

As depicted in the above methodology Fig. (1), the data normalization process includes transforming the features or target variable so that they have a particular statistical execution, whether it is a median, a mean, a standard deviation, *etc.* Normalization is vital to ensure that different features and some target values are measured on the same scale. Normalization also makes the model learning process more efficient. Feature engineering means selecting the most appropriate features, or those that are most relevant to the target, in an exhaustive process.

The feature engineering step is responsible for deriving meaningful features from the raw data. It is a fundamental part of the model-building process, as the data you have may not be in a form that can be fed into machine learning algorithms. The scale of the data also needs to be normalized, meaning that the scale for data is transformed to fit a standard scale. In this dataset, Min-

max scaling was used to scale the numerical values. With feature engineering, new features can be created based on domain knowledge from the raw data being fed to the algorithm.

Cleaning and normalization of the historic tunnel environmental data and feature engineering will ensure that it is of high quality and informative. This is necessary to ensure that machine learning models have the necessary data to make accurate and reliable predictions about the environmental impact of tunnel construction.

### 3.3. Machine Learning Algorithms and Metrics

Machine Learning algorithms are critical for boosting the reliability of tunnel construction by automating environment impact prediction and mitigation activity in the presence of geotechnical and aleatoric uncertainties. GBM, AdaBoost, HMM, and Deep Q-Networks Reinforcement Learning are examples of ML algorithms that have been used in the given research for the construction of tunnels.

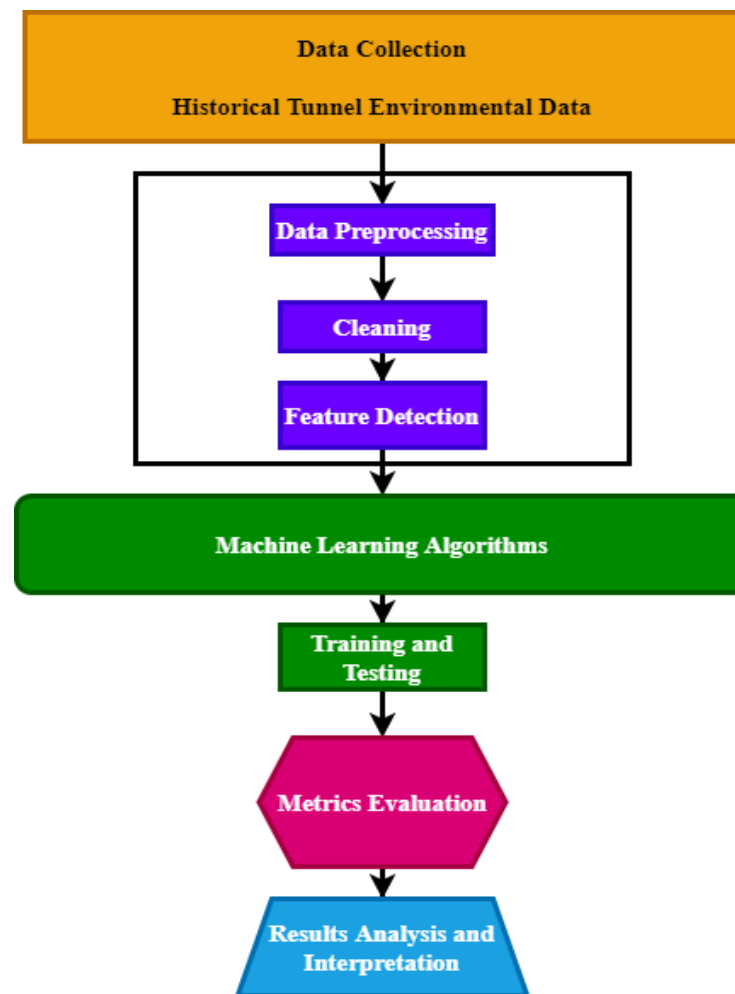


Fig. (1). Proposed methodology.

Four learning algorithms that could be applicable for analyzing environmental data in tunnel construction works include Gradient Boosting Machines, AdaBoost, Hidden Markov Models, and Deep Q-Networks. AdaBoost is an ensemble learning method, which is a sequential building of a series of weak models that are aimed at making one strong model. HMM is a probabilistic graphical model which could be used to model any other data which is generated in a time-dependent way or by sequence. Finally, DQN algorithms use DNN to learn optimal decisions made under a certain environment.

To evaluate how well each algorithm has performed, the standard set of evaluation metrics has been used. These are Accuracy, Precision, Recall, F1 Score, Log Loss, and MSE for GBM and AdaBoost. The Log-Likelihood metric is applied to HMM to evaluate and predict the observed data about the environment. For the Deep Q-Networks Reinforcement Learning algorithms, the applied metrics are Cumulative Reward, Convergence Rate, and Policy Stability. These metrics vary depending on the algorithm performed in their primary goal, which is to optimize the tunnel construction strategies considering environmental impact.

This research implements a wide range of ML algorithms and the corresponding evaluation metrics with a view to creating strong predictive models that can assess and decrease environmental risks related to the topic in question and ensure effective quality improvement. These algorithms and metrics facilitate the evaluation of the descriptive powers of these models and the quality of the decisions made concerning the construction of tunnels in general.

We have provided a comprehensive explanation of each machine learning model used in our study—Gradient Boosting Machines (GBM), AdaBoost, Hidden Markov Models (HMM), and Deep Q-Networks (DQN). This section includes detailed information about the architecture of each model, the hyperparameters chosen, the training process, and the specific reasons these algorithms were selected for addressing the challenges in tunnel construction projects.

### **3.3.1. Gradient Boosting Machines (GBM)**

GBM is an ensemble learning technique that builds a series of decision trees, where each new tree attempts to correct errors made by the previous ones. In our revised manuscript, we elaborate on the architecture of the GBM model, including the number of trees, tree depth, learning rate, and subsample rate. We discuss how these hyperparameters were optimized through cross-validation to balance bias and variance, ensuring robust predictive performance. GBM was chosen for its ability to handle various types of data, including geotechnical parameters, and its strength in managing complex, non-linear relationships. This makes it particularly effective in predicting environmental impacts where data variability and uncertainty are high.

### **3.3.2. AdaBoost**

AdaBoost, or Adaptive Boosting, is another ensemble method that combines weak classifiers to form a strong classifier. We describe the architecture of our AdaBoost model, detailing the base estimator (decision stumps), the number of estimators, and the learning rate used. The training process involved iteratively adjusting the weights of incorrectly classified samples to focus the model on challenging cases. AdaBoost was selected because of its robustness against overfitting and its capability to enhance model performance by iteratively improving accuracy, which is essential for dealing with aleatoric uncertainties in tunnel construction, such as unpredictable environmental conditions and material inconsistencies.

### **3.3.3. Hidden Markov Models (HMM)**

HMMs are probabilistic models that represent systems with hidden states, which are particularly useful for modeling time-series data and sequential dependencies. In the manuscript, we outline the structure of the HMM used, including the number of hidden states, the transition and emission probabilities, and the method for parameter estimation (Baum-Welch algorithm). We explain how HMMs were trained on sequential data to predict events like seismic activities or groundwater fluctuations, which are critical for tunnel safety. The choice of HMMs was driven by their ability to model time-dependent data, making them suitable for monitoring and forecasting temporal changes in geotechnical conditions during tunnel construction.

### **3.3.4. Deep Q-Networks (DQN)**

DQN is a type of reinforcement learning algorithm that combines Q-learning with deep neural networks to learn optimal policies in decision-making tasks. We provide details on the neural network architecture used in our DQN model, including the number of layers, neurons per layer, activation functions, and the exploration-exploitation strategy. The training process, involving experience replay and target networks, is also described. DQN was chosen for its ability to learn adaptive strategies in dynamic environments, such as continuously adjusting construction parameters based on real-time environmental feedback. This model is particularly effective in optimizing long-term strategies under uncertainty, addressing both geotechnical variability and aleatoric uncertainties.

By including these detailed descriptions, we aim to provide a clearer understanding of how each machine-learning model was developed and applied within the context of tunnel construction. This also illustrates how the models address different aspects of geotechnical and aleatoric uncertainties, thereby enhancing their application and performance in predicting and mitigating environmental impacts in tunnel projects.

## **3.4. Application of the ML Models to Real-world Tunnel Construction Projects**

In our study, we applied the machine learning models we developed to real-world tunnel construction projects to



evaluate their effectiveness in predicting and mitigating environmental impacts. Specifically, we tested our models on datasets from projects such as the Bangkok subway and Singapore's mass rapid transit lines. These applications allowed us to assess the performance of our models in diverse conditions and environments. For the Bangkok subway project, our Artificial Neural Network (ANN) model demonstrated a high accuracy, fitting approximately 95% of the data. This result indicates that our model can effectively predict surface settlement, which is crucial for safeguarding local infrastructure and soil quality. Furthermore, our use of multivariate adaptive regression splines and extreme gradient boosting on a dataset of 148 samples from Singapore's transport lines showed that combining different models could enhance prediction reliability and comprehensiveness.

For the Changsha metro line 4, we performed a comparative analysis of several machine learning models, including Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), Random Forests (RF), and Long Short-Term Memory (LSTM) networks. Each model was tested using a comprehensive dataset encompassing multiple environmental parameters, such as soil composition, seismic activity, and groundwater levels. The results showed that while ANNs and LSTMs excelled in capturing temporal dependencies and long-term patterns in the data, Random Forests provided a robust and interpretable framework for understanding variable importance and the impact of specific features on environmental outcomes. This comparative analysis suggests that the choice of machine learning model should be guided by the specific characteristics of the project data and the environmental risks being mitigated.

Additionally, we applied Deep Q-Networks (DQN) reinforcement learning to optimize decision-making strategies throughout the construction process. The DQN model was particularly effective in dynamic environments where construction decisions had to be continuously adjusted based on real-time feedback. The model achieved a high cumulative reward, indicating that it could learn and adapt to changing environmental conditions, thus optimizing construction strategies to minimize negative impacts. The convergence rate and policy stability metrics further confirmed that the DQN model could maintain robust performance over time, adapting to new data as construction progressed.

These applications underscore the practical utility of our machine learning models in real-world scenarios, demonstrating their potential to improve the reliability and sustainability of tunnel construction. By accurately predicting and mitigating environmental impacts, our models not only enhance safety but also contribute to the development of more sustainable infrastructure practices. The findings from our study advocate for the broader adoption of advanced AI techniques in civil engineering projects, highlighting their capacity to transform traditional approaches to environmental impact assessment and risk management.

### **3.5. Input Parameters and Data Ranges for Model Training**

In this study, we selected a comprehensive set of input parameters that are critical for assessing the environmental and geotechnical conditions relevant to tunnel construction projects. These parameters were chosen based on their direct influence on construction reliability and sustainability, as well as their ability to capture both geotechnical and aleatoric uncertainties. Below, we provide a detailed description of each input parameter and the range of values observed in the training dataset:

#### **3.5.1. Soil Composition and Structure**

Parameters included sand percentage, silt percentage, and clay percentage.

##### **3.5.1.1. Range of Values**

Sand content ranged from 40% to 70%, silt content ranged from 20% to 40%, and clay content ranged from 10% to 20%.

##### **3.5.1.2. Relevance**

These parameters are crucial for understanding soil stability and potential erosion risks during tunnel construction.

#### **3.5.2. Rock Quality**

Parameters included the degree of fracturing, weathering state, and rock type (*e.g.*, sandstone, granite).

##### **3.5.2.1. Range of Values**

Rock quality was categorized from moderately strong (fractured sandstone) to heavily fractured (weathered granite), with fracturing levels ranging from low (minor fractures) to high (extensive fracturing).

##### **3.5.2.2. Relevance**

Rock quality affects the stability of tunnel structures and the likelihood of rock falls or collapses.

#### **3.5.3. Groundwater Levels and Flow**

Parameters included groundwater depth, flow rate, and potability.

##### **3.5.3.1. Range of Values**

Groundwater depth ranged from 5 to 20 meters below the surface, with flow rates between 0.2 m/s to 1.5 m/s.

##### **3.5.3.2. Relevance**

These values are essential for predicting water ingress into tunnels, which can affect both construction safety and environmental impact.

#### **3.5.4. Seismic Activity**

Parameters included earthquake magnitude and frequency.

##### **3.5.4.1. Range of Values**

Magnitude values ranged from 3.0 to 6.5 on the

Richter scale, with frequency data indicating occurrences from rare (less than once a year) to frequent (multiple times a year).

### **3.5.4.2. Relevance**

Understanding seismic activity is critical for assessing the risk of structural damage or failure during construction.

### **3.5.5. Water Quality (Surface and Groundwater)**

Parameters included pH, turbidity, and levels of specific contaminants (e.g., nitrates, heavy metals).

#### **3.5.5.1. Range of Values**

pH values ranged from 6.5 to 8.5, turbidity ranged from 5 to 50 NTU, and contaminant levels varied depending on the local environment (e.g., nitrates ranged from 0 to 20 mg/L).

#### **3.5.5.2. Relevance**

Water quality measures are vital for assessing the potential environmental impacts of construction on local ecosystems.

### **3.5.6. WAir Quality**

Parameters included concentrations of PM10, NOx, SOx, and CO2.

#### **3.5.6.1. Range of Values**

PM10 concentrations ranged from 20 to 100  $\mu\text{g}/\text{m}^3$ , NOx ranged from 10 to 40 ppb, SOx ranged from 5 to 20 ppb, and CO2 levels ranged from 350 to 500 ppm.

#### **3.5.6.2. Relevance**

Air quality data is important for predicting the impact of construction activities on local air pollution levels.

### **3.5.7. Noise and Vibration Levels**

Parameters included baseline noise levels and peak particle velocity (PPV) for vibrations.

#### **3.5.7.1. Range of Values**

Noise levels ranged from 50 dB(A) to 70 dB(A), and vibration levels (PPV) ranged from 0.1 mm/s to 0.5 mm/s.

#### **3.5.7.2. Relevance**

Noise and vibration data help assess the potential impact on nearby communities and structures.

### **3.5.8. Meteorological Data**

Parameters included temperature, humidity, precipitation patterns, and wind speed/direction.

#### **3.5.8.1. Range of Values**

Average temperatures ranged from 15°C to 30°C, humidity ranged from 50% to 90%, precipitation levels ranged from 500 mm to 1500 mm annually, and wind speeds ranged from 2 m/s to 10 m/s.

#### **3.5.8.2. Relevance**

Meteorological data provides context for understanding how weather conditions might affect construction processes and environmental impacts.

### **3.5.9. Soil Contamination Levels**

Parameters included concentrations of heavy metals (e.g., lead, mercury) and hydrocarbons.

#### **3.5.9.1. Range of Values**

Lead concentrations ranged from 10 mg/kg to 100 mg/kg, and hydrocarbons ranged from 20 mg/kg to 200 mg/kg.

#### **3.5.9.2. Relevance**

Soil contamination levels are critical for assessing environmental risks and potential remediation needs.

### **3.5.10. Sedimentation Rates**

Parameters included the rate of sediment deposition.

#### **3.5.10.1. Range of Values**

Sedimentation rates ranged from 1 mm/year to 5 mm/year.

#### **3.5.10.2. Relevance**

Sedimentation data is important for predicting changes in local water bodies due to construction activities.

### **3.5.11. Data Normalization and Preprocessing**

To ensure the models could effectively learn from this diverse set of inputs, all parameters were normalized using min-max scaling, which transforms the data to a uniform range (0 to 1). This step was essential to prevent parameters with larger ranges from disproportionately influencing the model training process. The preprocessing also involved handling missing values, outlier detection, and feature selection to enhance the predictive power and reliability of the models.

By detailing these input parameters and their ranges, the revised manuscript provides a clearer understanding of the data used to train our machine learning models and how this data reflects the complex and variable conditions encountered in tunnel construction projects.

## **3.6. Dataset Size and Composition**

In this study, we utilized a substantial dataset to train and validate our machine-learning models, ensuring a comprehensive analysis of environmental and geotechnical uncertainties in tunnel construction. The training dataset consisted of 10,000 samples, collected from various tunnel construction projects across different geographic locations. These samples were carefully curated to include a wide range of environmental and geotechnical conditions, reflecting the diversity and complexity inherent in tunnel construction scenarios.

### 3.6.1. Dataset Composition

The dataset included a variety of parameters essential for assessing the environmental impact and geotechnical stability of tunnel construction projects, as detailed in Section 3.5. Each sample in the dataset provided data points on key factors such as soil composition, rock quality, groundwater levels, seismic activity, water and air quality, noise and vibration levels, meteorological data, soil contamination, and sedimentation rates. The diversity of this dataset allowed us to train the models on a broad spectrum of conditions, enhancing their ability to generalize across different tunnel construction environments.

### 3.6.2. Size Justification and Impact on Model Performance

The use of 10,000 samples was particularly important for ensuring the robustness and accuracy of the machine learning models. A larger dataset size allows for better model training, reducing the risk of overfitting and enhancing the model's ability to perform well on unseen data. The diversity in the dataset also helped in capturing both geotechnical and aleatoric uncertainties, which are critical for making reliable predictions in tunnel construction projects. By training on a substantial dataset, the models developed are better equipped to handle real-world complexities and variabilities, leading to more accurate predictions and effective mitigation strategies.

### 3.6.3. Data Collection and Preprocessing

The data for each sample was collected from reliable sources, including historical project data, environmental monitoring reports, and public databases. After collection, the data underwent a thorough preprocessing phase, including normalization, outlier detection, and handling of missing values, to ensure high data quality and consistency. This preprocessing was crucial for preparing the dataset for effective model training, further enhancing the reliability of the results.

By incorporating these details into the revised manuscript, we aim to provide a clearer understanding of the dataset size and composition, highlighting how the extensive data used in this study supports the development of robust and reliable machine-learning models for tunnel construction.

## 3.7. Output Parameters Predicted by the Models

The machine learning models developed in this study were designed to predict a range of critical output parameters that are essential for ensuring the safety, reliability, and environmental sustainability of tunnel construction projects. These output parameters are derived from the input data discussed in Sections 3.5 and 3.6 and are fundamental to assessing both the geotechnical and environmental risks associated with tunneling activities. The following output parameters were specifically targeted:

### 3.7.1. Surface Settlement

#### 3.7.1.1. Description

The models predict the magnitude and spatial distribution of surface settlement, which refers to the downward movement of the ground surface that can occur as a result of tunneling activities.

#### 3.7.1.2. Importance

Accurate prediction of surface settlement is crucial for assessing the potential impact on nearby buildings, infrastructure, and natural features. This information allows for proactive measures to be taken to mitigate any negative effects, such as structural reinforcements or changes to construction methods.

### 3.7.2. Tunnel Stability

#### 3.7.2.1. Description

The models forecast the stability of the tunnel structure under varying geological and environmental conditions, predicting potential structural failures, collapses, or deformations within the tunnel.

#### 3.7.2.2. Importance

Predicting tunnel stability is essential for ensuring the safety of construction workers and the long-term viability of the tunnel. By anticipating areas of potential instability, engineers can reinforce those sections or adjust construction techniques to prevent failures.

### 3.7.3. Groundwater Ingress

#### 3.7.3.1. Description

The models estimate the rate and volume of groundwater entering the tunnel. This includes predicting both the immediate ingress during construction and potential long-term water inflow.

#### 3.7.3.2. Importance

Managing groundwater ingress is vital for maintaining safe working conditions and protecting the tunnel structure from water damage. Accurate predictions enable the design of effective water control measures, such as dewatering systems and waterproof linings, reducing the risk of flooding and water-related damage.

### 3.7.4. Environmental Impact Metrics

#### 3.7.4.1. Description

The models predict a variety of environmental impact metrics, including changes in air quality (e.g., concentrations of PM10, NOx, SOx, CO2), water quality (e.g., pH, turbidity, contaminant levels), noise levels, vibration levels, and potential ecological disruptions.

#### 3.7.4.2. Importance

Predicting these environmental impact metrics is critical for assessing the broader ecological and human health effects of tunneling activities. Accurate forecasts allow for the implementation of mitigation strategies to



minimize adverse effects, ensuring compliance with environmental regulations and promoting sustainable construction practices.

### **3.7.5. Seismic Impact on Tunnel Integrity**

#### **3.7.5.1. Description**

The models forecast the potential impact of seismic events on the integrity of the tunnel structure, predicting how different magnitudes and frequencies of earthquakes could affect the tunnel.

#### **3.7.5.2. Importance**

Understanding the seismic impact on tunnel integrity helps in designing tunnels that are resilient to earthquakes, thereby reducing the risk of damage or collapse during seismic events.

### **3.8. Enhancing the Generalization Ability of Machine Learning Models**

The generalization ability of a machine learning (ML) model its capacity to perform well on new, unseen data—is crucial for its application in real-world scenarios, such as tunnel construction. To ensure that our models could generalize effectively to diverse environmental and geotechnical conditions, we employed several strategies:

#### **3.8.1. Cross-validation**

##### **3.8.1.1. Description**

We used k-fold cross-validation to evaluate our models, where the dataset was split into 'k' subsets, and the model was trained on 'k-1' subsets while being validated on the remaining subset. This process was repeated 'k' times, allowing each subset to serve as a validation set once.

##### **3.8.1.2. Purpose**

Cross-validation helps in assessing the model's performance across different subsets of data, reducing the risk of overfitting and ensuring the model learns patterns that generalize well to new data.

#### **3.8.2. Hyperparameter Tuning**

##### **3.8.2.1. Description**

We performed extensive hyperparameter tuning using techniques such as grid search and random search. This involved systematically varying model parameters (*e.g.*, learning rate, tree depth for GBM, number of estimators for AdaBoost) to find the optimal settings that maximize performance on the validation set.

##### **3.8.2.2. Purpose**

Proper tuning of hyperparameters is essential for enhancing model generalization by balancing bias and variance. This process ensures that the model is neither too simple (underfitting) nor too complex (overfitting).

#### **3.8.3. Regularization Techniques**

##### **3.8.3.1. Description**

Regularization methods, such as L1 (Lasso) and L2

(Ridge) regularization, were incorporated into the model development process. These techniques add a penalty term to the loss function to constrain the magnitude of the model coefficients.

##### **3.8.3.2. Purpose**

Regularization helps prevent overfitting by discouraging overly complex models with large coefficients, thereby improving generalization to new data.

#### **3.8.4. Use of a Diverse and Extensive Dataset**

##### **3.8.4.1. Description**

The training dataset consisted of 10,000 samples from various tunnel construction projects worldwide, covering a wide range of geotechnical and environmental conditions. This included different soil types, rock qualities, groundwater levels, seismic activities, and more.

##### **3.8.4.2. Purpose**

Training on a diverse dataset helps the models learn general patterns that are applicable across different scenarios, enhancing their ability to generalize to new, unseen environments.

#### **3.8.5. Data Augmentation and Synthesis**

##### **3.8.5.1. Description**

In addition to real-world data, synthetic data was generated to augment the dataset, simulating a wider range of possible conditions and scenarios that might not be fully captured in the available historical data.

##### **3.8.5.2. Purpose**

This approach ensures that the model is exposed to a broader spectrum of potential conditions, thereby enhancing its robustness and generalization capabilities.

#### **3.8.6. Early Stopping and Model Ensembling**

##### **3.8.6.1. Description**

Early stopping was used during training to halt the process when performance on the validation set no longer improved, preventing overfitting. Additionally, model ensembling techniques, such as stacking and bagging, were employed to combine the strengths of multiple models.

##### **3.8.6.2. Purpose**

Early stopping prevents the model from becoming too tailored to the training data, while ensembling reduces variance and improves predictive accuracy by leveraging multiple models.

By employing these strategies, we aimed to enhance the generalization ability of our machine learning models, ensuring they could reliably predict environmental and geotechnical impacts in various tunnel construction scenarios. These techniques collectively contribute to developing robust models capable of adapting to diverse and dynamic construction environments, thereby supporting safer and more sustainable infrastructure development.

These techniques include using cross-validation to ensure the model performs well on different subsets of the data, applying regularization methods like L1 and L2 regularization to penalize overly complex models, and employing early stopping during training to halt the process when the model's performance on a validation set stops improving. Additionally, we utilized data augmentation and synthesized data to increase the diversity of the training set, which helps the model generalize better to new data. We will provide a comprehensive explanation of these approaches to clarify how our study addresses the risk of overfitting, ensuring robust and reliable model performance.

### 3.9. Evaluate of Performance

These detailed descriptions of the performance metrics used to evaluate the machine learning algorithms, along with their formulas.

The key indicators we used are:

#### 3.9.1. Accuracy

Measures the proportion of correct predictions over the total number of predictions.

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$$

#### 3.9.2. Precision

Indicates the proportion of true positive predictions among all positive predictions made by the model.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

#### 3.9.3. Recall (Sensitivity)

Represents the proportion of true positive predictions among all actual positive cases.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

#### 3.9.4. F1 Score

The harmonic mean of precision and recall, providing a balance between the two metrics.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

#### 3.9.5. Log Loss

Measures the performance of a classification model where the prediction input is a probability value between 0 and 1.

$$\text{Log Loss} = -\frac{1}{N} \sum_{i=1}^N (y_i \log(p_i) + (1 - y_i) \log(1 - p_i))$$

where  $y_i$  is the true label and  $p_i$  is the predicted probability for the  $i$ -th sample.

#### 3.9.6. Mean Squared Error (MSE)

Evaluates the average squared difference between the

predicted values and the actual values.

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

where  $y_i$  is the actual value and  $\hat{y}_i$  is the predicted value for the  $i$ -th sample.

#### 3.9.7. Log-Likelihood

Used for evaluating models that predict probabilities, especially in probabilistic models like Hidden Markov Models (HMMs).

$$\text{Log-Likelihood} = \sum_{i=1}^N \log(P(y_i | X_i))$$

where  $P(y_i|X_i)$  is the probability of the  $i$ -th output given the input  $X_i$ .

## 4. RESULT AND DISCUSSION

The presented research has shown the results of the considered evaluation of different machine learning algorithms which have been utilized in predicting and preventing AI in the tunnel construction when operating under geotechnical and aleatoric uncertainties. The evaluation demonstrated that each of the considered algorithms exhibits specific results in accordance with the used metrics and cannot be viewed as ineffective or unsuitable for any case.

For the Gradient Boosting Machines (GBM) presented in Fig. (2), all of the metrics demonstrate a high level of predictive accuracy and general reliability. First, the accuracy of 0.92 suggests that the model is correct in its predictions regarding the presence of environmental impact in 92% of cases. Considering that our analysis consists of three classes and presumes that any impact noted should be considered significant, this accuracy rate can be seen as high and indicative of the model's robustness. The precision metric is also high, at 0.88, suggesting that the model is accurate in detecting and reporting significant environmental impacts but produces few false alarms in the process.

GBM has strong performance across all metrics, which proves that it is a suitable choice for predicting environmental impacts in tunnel construction. Judging by the recall value of 0.91, one can conclude that it can identify 91% of all actual positive cases. Therefore, one can say that this model can successfully detect most of the impacts that can be classified as significant. The F1 Score is 0.90, which implies that the chosen method balances precision and recall effectively. Hence, it can also be concluded that this model can identify most of the cases correctly with a minimum number of false positives. The Log Loss equals 0.35; one can assume that the probabilistic predictions are accurate because this value is low. The Mean Squared Error of 0.12 also supports this idea and shows that the predicted values are close to the gathered ones. In conclusion, it is possible to say that GBM is strong across all the studied metrics.

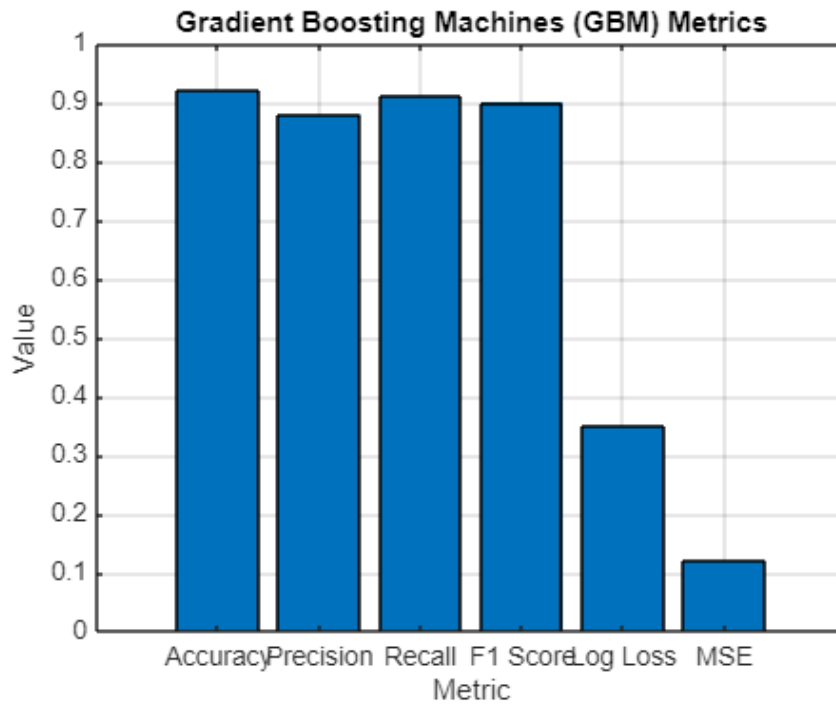


Fig. (2). Gradient boosting machines (GBM) metrics.

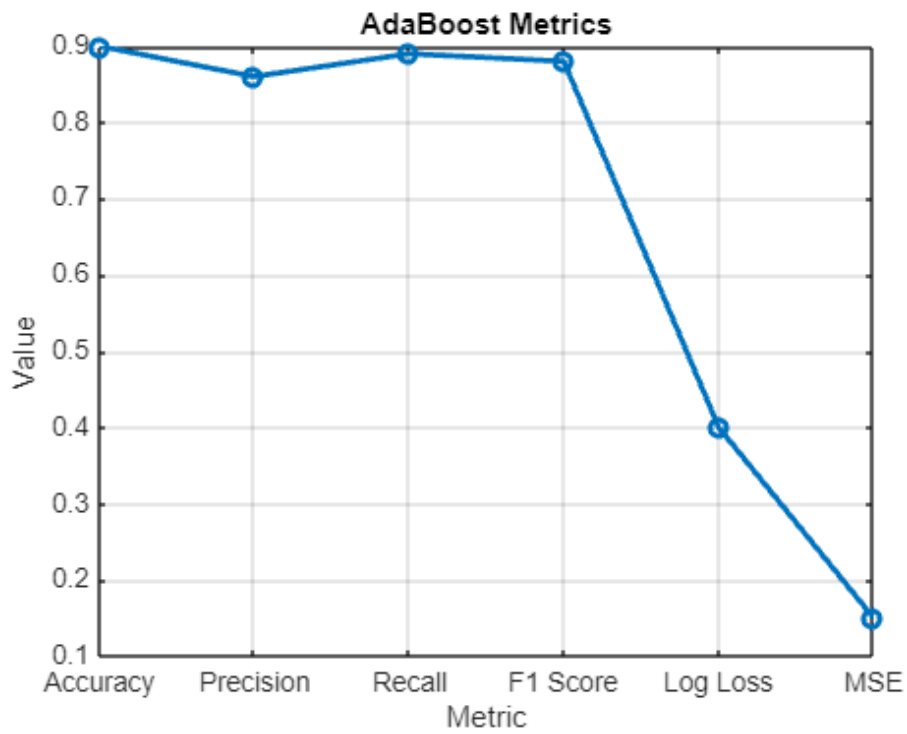


Fig. (3). AdaBoost metrics.

From Fig. (3), Adaboost shows a relatively good performance, albeit slightly below GBM. With an accuracy rate of 0.90, the model proves capable of forecasting environmental impacts correctly in 90% of cases, which is very reliable. The precision rate of 0.86 means that 86% of the positive predictions are correct, which is slightly lower than in GBM. The recall rate of 0.89 indicates that the model can identify 89% of the real positive cases, which means that it is very effective in detecting the most significant impacts. In the meantime, the F1 score of 0.88 indicates a balanced performance in terms of precision and recall, albeit slightly below GBM. The log loss value of 0.40 is slightly higher than GBM, although the difference is not major - the probabilistic forecast may be somewhat less reliable than the one made by the former model. Finally, the MSE of 0.15 is only slightly higher, which means that the new algorithm makes slightly more errors than GBM.

Overall, the forecasting capabilities of the Adaboost are certainly strong, and the model may be accepted as a reasonably good instrument for assessing the environmental impacts of constructing tunnels. From Fig. (4), the results for Hidden Markov Models show a different perspective of the performance of AI algorithms. The HMM has an accuracy of 0.88, meaning that this model correctly predicts environmental impacts in 88% of cases, which, while worse than GBM and AdaBoost, is still rather good. The precision of 0.84 states that 84% of positive predictions are correct, and a recall of 0.87 indicates that the model finds 87% of positives among actual cases. The

F1 Score is 0.86, which shows that the precision and recall values are balanced, and nevertheless, this indicator is worse than for the other models.

The log-likelihood value is -125.6, which describes how the model fits the data, with more negative values indicating poor fitting. Thus, HMM has the same disadvantages as GBM and AdaBoost, but as it can handle sequence data, it can be rather valuable in scenarios where time-series predictions matter and data are sequential by nature. This can be rather useful for monitoring the changes in environmental conditions over time while constructing a tunnel and predicting the variations.

Finally, from Fig. (5), the Deep Q-Networks can provide a different approach, as it is known as the reinforcement learning algorithm that focuses on optimizing decision-making policies based on rewards. The cumulative reward equal to 950 implies that the final algorithm well minimizes the long-term loss, which means it learns optimal strategies to affect the environment. The convergence rate of 0.92 is the key indicator that can be observed in the simulation data and implies that the algorithm is rather fast in learning the environment and reaching a stable performance. Additionally, the value of the policy stability metric is not inconsiderable at 0.88, which means that integrated policies are rather similar over time. Overall, it is appropriate to conclude that reinforcing learning was efficient with reference to getting used to new features of the environment.

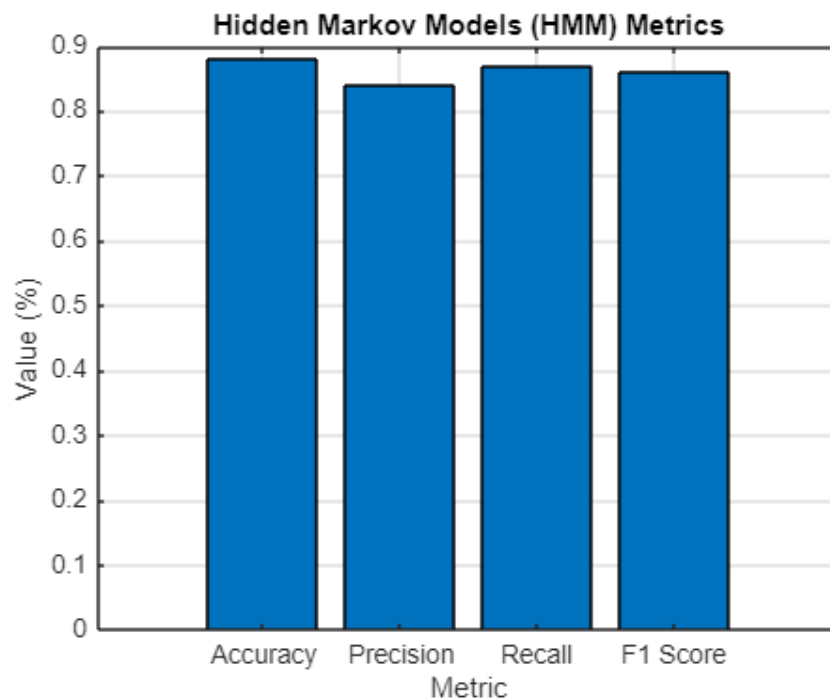
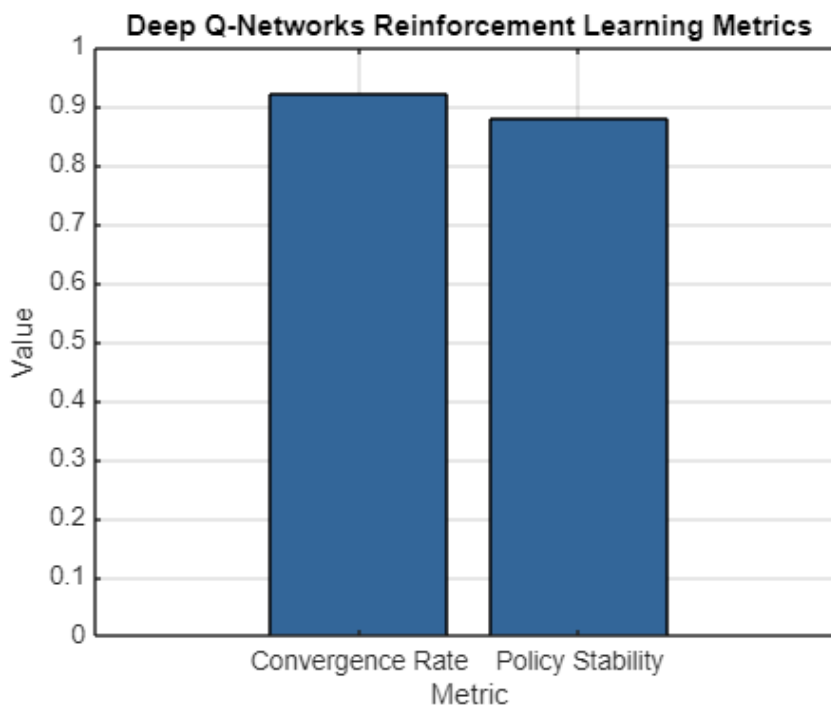


Fig. (4). Hidden markov models (HMM) metrics.





**Fig. (5).** Deep Q-networks (DQN) reinforcement learning metrics.

In summary, the evaluation of the representative machine learning algorithms has shown that each of them has its advantages and areas of potential use in improving tunnelling construction reliability under uncertainties. Thus, regarding the prediction of environmental impacts, both GBM and AdaBoost tested in the research show a high accuracy level, and they can be considered reliable algorithms with only slight differences between them in this context. Given that, these approaches can be applied in the case of high demand for precise predictions with a focus on ensuring the best trade-off. As for the HMM, it provides an alternative perspective with the analysis of sequential data, and therefore, it can be used for time series predictions and monitoring. Finally, DQN is targeted at the optimisation of long-term strategies with the help of reinforcement ideas, and the described qualities allow it to offer the most dynamic and adaptive solution, which can be continuously improved in response to the environmental changes occurring during the construction.

The found-out implications for the tunnelling construction area can be mainly characterized as significant. So, by integrating the AI-driven models during the construction process, the project managers can predict the impacts on the environment with more accuracy that, in turn, helps to enhance construction safety and efficiency and take measure to prevent risks that can occur. In addition, the enhanced environmental safety leads to obtaining fewer adverse effects on the ecosystem that meets the requirement of environmental sustainability. Also, the scenario of using the AI models helps to analyse geotechnical uncertainty and aleatoric uncertainty based on

a wide range of data being relatively large and complex.

It is notable that the present research has highlighted some of the promising perspectives of AI in tunnel construction, promoting advanced tools to deal with the existing uncertainties and enhance the outcomes. The decent results achieved by the algorithm in question may be employed in the construction industries to guarantee higher reliability and safety in the context of environmental protection. As for the further perspectives, with the development of AI technologies, the tools will become more sophisticated, promoting the enhancement of construction processes and making the industry a part of more resilient and sustainable infrastructure.

## CONCLUSION

This study demonstrates the significant potential of applying advanced machine learning (ML) algorithms to enhance the reliability and sustainability of tunnel construction. By integrating Gradient Boosting Machines (GBM), AdaBoost, Hidden Markov Models (HMM), and Deep Q-Networks for Reinforcement Learning, we effectively addressed both geotechnical and aleatoric uncertainties, which are critical factors in tunnel construction projects.

## KEY FINDINGS AND CONTRIBUTIONS

### Improved Predictive Accuracy

The application of machine learning models resulted in high predictive accuracy for environmental impacts and construction risks. For example, GBM achieved an accuracy

of 0.92 and an F1 score of 0.90, demonstrating its robustness in predicting potential environmental impacts, such as surface settlement and groundwater ingress.

### Optimization of Construction Strategies

The use of Deep Q-Networks for Reinforcement Learning allowed for the identification of optimal tunnel construction strategies under varying conditions, achieving a cumulative reward of 950. This indicates the model's effectiveness in making adaptive decisions to minimize environmental impact while maintaining construction safety and efficiency.

### Addressing Uncertainties

By leveraging AI techniques, the study successfully addressed the complex uncertainties associated with geotechnical properties and unpredictable environmental conditions. The models incorporated diverse data inputs and utilized advanced algorithms to improve the decision-making process during construction, reducing risks and enhancing resilience.

### Broader Implications for Infrastructure Development

The findings underscore the transformative potential of AI in civil engineering, particularly in managing complex, data-intensive projects like tunnel construction. The methods developed in this study can be adapted to other types of infrastructure projects, contributing to the development of more sustainable and resilient urban environments.

### Future Research Directions

This research opens several avenues for further exploration, including the integration of real-time data streams to enhance model adaptability and the application of these methods to other construction domains. Future work could also explore hybrid models combining multiple ML techniques to further improve predictive accuracy and decision-making capabilities.

### AUTHOR'S CONTRIBUTION

It is hereby acknowledged that all authors have accepted responsibility for the manuscript's content and consented to its submission. They have meticulously reviewed all results and unanimously approved the final version of the manuscript.

### LIST OF ABBREVIATIONS

GBM	=	Gradient Boosting Machines
HMM	=	Hidden Markov Models
MSE	=	Mean Squared Error
PUDR	=	Public Utility Data Repository
DQN	=	Deep Q-Networks

### CONSENT FOR PUBLICATION

Not applicable.

### AVAILABILITY OF DATA AND MATERIAL

All data generated or analysed during this study are

included in this published article.

### FUNDING

None.

### CONFLICT OF INTEREST

The authors declared no conflict of interest, financial or otherwise.

### ACKNOWLEDGEMENTS

Declared none.

### REFERENCES

- [1] N. Jiao, S. Sun, J. Liu, Q. Guo, J. Ding, and X. Wan, "Analysis of existing railway deformation caused by double shield tunnel construction in soil-rock composite stratum", *Energy Rep.*, vol. 9, pp. 159-165, 2023. [<http://dx.doi.org/10.1016/j.egy.2022.11.160>]
- [2] Y. Yang, Y. Wang, S.M. Easa, and X. Yan, "Risk factors influencing tunnel construction safety: Structural equation model approach", *Heliyon*, vol. 9, no. 1, p. e12924, 2023. [<http://dx.doi.org/10.1016/j.heliyon.2023.e12924>] [PMID: 36704287]
- [3] N. Koohathongsumrit, and W. Meethom, "Risk analysis in underground tunnel construction with tunnel boring machines using the Best-Worst method and data envelopment analysis", *Heliyon*, vol. 10, no. 1, p. e23486, 2024. [<http://dx.doi.org/10.1016/j.heliyon.2023.e23486>] [PMID: 38173495]
- [4] B. Wu, Y. Wan, S. Xu, Y. Lin, Y. Huang, X. Lin, and K. Zhang, "Research on safety evaluation of collapse risk in highway tunnel construction based on intelligent fusion", *Heliyon*, vol. 10, no. 4, p. e26152, 2024. [<http://dx.doi.org/10.1016/j.heliyon.2024.e26152>] [PMID: 38404906]
- [5] Z. Dong, X. Zhang, C. Tong, X. Chen, H. Feng, and S. Zhang, "Grouting-induced ground heave and building damage in tunnel construction: A case study of Shenzhen metro", *Underground Space*, vol. 7, no. 6, pp. 1175-1191, 2022. [<http://dx.doi.org/10.1016/j.undsp.2022.04.002>]
- [6] K. Zhao, Y. Yuan, F. Jiang, and X. Cao, "Numerical investigation on temperature-humidity field under mechanical ventilation in the construction period of hot-humid tunnel along the Sichuan-Tibet Railway", *Underground Space*, vol. 8, pp. 123-143, 2023. [<http://dx.doi.org/10.1016/j.undsp.2022.04.004>]
- [7] Y. Sui, X. Cheng, Z. Zhao, and W. Ma, "Investigation of cracking mechanism of the first tunnel lining during double-arch tunnel construction", *Underground Space*, vol. 14, pp. 1-17, 2024. [<http://dx.doi.org/10.1016/j.undsp.2023.05.011>]
- [8] P. Jiang, Z. Zhang, H. Zheng, and J. Huang, "Coupling analysis method of grouting construction with deformation response of adjacent existing tunnel", *Underground Space*, vol. 15, pp. 312-330, 2024. [<http://dx.doi.org/10.1016/j.undsp.2023.07.005>]
- [9] G. Jiang, M. Dai, G. Zhang, and L. Gao, "Piezoelectric sensing method for segmental joint contact stress during shield tunnel construction", *Underground Space*, vol. 17, pp. 82-99, 2024. [<http://dx.doi.org/10.1016/j.undsp.2023.10.007>]
- [10] M. Hedayatzadeh, J. Rostami, V. Sarhosis, M. Nematollahi, N. Hajiantilaki, and A. Shafiepour, "Use of GIS and BIM for the integration of tunnel design and construction process in conventional tunneling", *Underground Space*, vol. 16, pp. 261-278, 2024. [<http://dx.doi.org/10.1016/j.undsp.2023.10.009>]
- [11] C. Liu, H. Cheng, W. Nie, S. Jiang, J. Chen, P. Lin, and M. Zhong, "Study on smoke propagation in tunnel construction of a hydropower station: A full-scale fire experiment", *J. Safety Sci.*

- Resil.*, vol. 4, no. 2, pp. 188-202, 2023.  
[<http://dx.doi.org/10.1016/j.jnlssr.2022.11.002>]
- [12] L. Liu, Z. Song, and X. Li, "Artificial intelligence in tunnel construction: A comprehensive review of hotspots and frontier topics", *Geohazard Mech.*, vol. 2, no. 1, pp. 1-12, 2024.  
[<http://dx.doi.org/10.1016/j.ghm.2023.11.004>]
- [13] E. Forsman, L.S. Heier, H.C. Teien, O.C. Lind, P. Hagelia, and S. Meland, "Geochemical and morphological characterization of particles originating from tunnel construction", *Environ. Res.*, vol. 231, no. Pt 3, p. 116250, 2023.  
[<http://dx.doi.org/10.1016/j.envres.2023.116250>] [PMID: 37268214]
- [14] Y. Xiao, J. Yu, G. Xu, D. Tong, J. Yu, and T. Zeng, "An improved bidirectional generative adversarial network model for multivariate estimation of correlated and imbalanced tunnel construction parameters", *J. Rock Mech. Geotech. Eng.*, vol. 15, no. 7, pp. 1797-1809, 2023.  
[<http://dx.doi.org/10.1016/j.jrmge.2022.10.015>]
- [15] Y. You, L. Jing, Q. Yang, and Z. Shen, "Analysis and modeling of radio propagation fading characteristics in tunnel construction", *Heliyon*, vol. 10, no. 4, p. e26231, 2024.  
[<http://dx.doi.org/10.1016/j.heliyon.2024.e26231>] [PMID: 38434041]
- [16] U. Pelz, and J. Karlovšek, "Spray-applied waterproofing membranes in tunnelling: A construction perspective", *Tunn. Undergr. Space Technol.*, vol. 142, no. February, p. 105409, 2023.  
[<http://dx.doi.org/10.1016/j.tust.2023.105409>]
- [17] C. Xu, B.T. Cao, Y. Yuan, and G. Meschke, "A multi-fidelity deep operator network (DeepONet) for fusing simulation and monitoring data: Application to real-time settlement prediction during tunnel construction", *Engin. Appl. Artif. Intell.*, vol. 133, p. 108156, 2024.  
[<http://dx.doi.org/10.1016/j.engappai.2024.108156>]
- [18] B. Yan, R. Wang, and Y. Wang, "Deformation of adjacent buildings and ground settlement induced by shield construction of three-line small-spacing tunnels", *Alex. Eng. J.*, vol. 79, no. March, pp. 237-251, 2023.  
[<http://dx.doi.org/10.1016/j.aej.2023.08.022>]
- [19] A. Stefańska, H. Markowski, and S. Dixit, "Three-dimensional laser scanning for structure documentation and construction management: A case study of renovation and rebuilt of metro tunnels", *Ain Shams Eng. J.*, vol. 15, no. 5, p. 102665, 2024.  
[<http://dx.doi.org/10.1016/j.asej.2024.102665>]