applied sciences

MDPI

Article

Digital Twin-Driven Framework for TBM Performance Prediction, Visualization, and Monitoring through Machine Learning

Kamran Latif [1], Abubakar Sharafat [2] and Jongwon Seo [1,*]

[1] Department of Civil & Environmental Engineering, Hanyang University, Seoul 04763, Republic of Korea; kamran5latif@gmail.com
[2] School of Architectural, Civil, Environment and Energy Engineering, Kyungpook National University, Daegu 41566, Republic of Korea; abubakarsharafat89@gmail.com
* Correspondence: jseo@hanyang.ac.kr

Abstract: The rapid development in underground infrastructure is encouraging faster and more modern ways, such as TBM tunneling, to meet the needs of the world. However, tunneling activities generate complex and heterogeneous data, which makes it difficult to visualize the performance of a project. Advancements in information technology, such as digital twins and machine learning, provide platforms for digital demonstration, visualization, and system performance monitoring of such data. Therefore, this study proposes a digital twin-driven framework for TBM performance prediction through machine learning, visualization, and monitoring. This novel approach integrates machine learning and real-time performance data to predict, visualize, and monitor the status of the tunnel construction progress. A digital twin virtual model of TBM was constructed based on TBM design parameters, the input parameter, boring energy, RPM, torque, thrust force, speed, gripper pressure, total revolution, and Q-value provided to SVR and ANN models to predict the TBM AR and PR, and TBM daily progress was visualized continuously. The predictive performance indices R^2 (0.97) and RMSE (0.011) were estimated for AR prediction, showing the accuracy of the proposed model. To demonstrate the proposed framework, this study shows the its effectiveness. By implementing this framework, stakeholders can minimize the risk associated with the cost and schedule of a tunneling project by simultaneously visualizing and monitoring the performance of TBMs through digital twin and machine learning algorithms.

Keywords: tunnel boring machine (TBM); TBM performance; digital twin (DT); machine learning (ML); visualization; monitoring

check for updates

Citation: Latif, K.; Sharafat, A.; Seo, J. Digital Twin-Driven Framework for TBM Performance Prediction, Visualization, and Monitoring through Machine Learning. *Appl. Sci.* **2023**, *13*, 11435. https://doi.org/10.3390/app132011435

Academic Editor: Tiago Miranda

Received: 30 September 2023
Revised: 13 October 2023
Accepted: 16 October 2023
Published: 18 October 2023

1. Introduction

In the current era of urbanization and space-limited environments, tunnels have become important infrastructure. They provide ease in transportation, water conveyance, storage, mining, defense facilities, dams, and flood control projects [1]. Modernization of tunnel excavation tools, from bone, gunpowder, and dynamite to mechanized tunneling, especially tunnel boring machines (TBMs), has revolutionized the construction industry. Various types of TBMs for different geological conditions and sizes are currently used in the construction industry, such as in subways, railways, water conveyance, and mining projects [2]. A TBM's application in long tunnels provides a fast, safe, and economical construction process [3]. However, TBMs are very sensitive to the geotechnical condition of the host rock mass throughout the tunnel [4]. The geological, operational, physical, and financial risks associated with tunneling make it a high-risk construction industry [5,6]. Penetration rate (PR), utilization index (UI), and advance rate (AR) are the key factors of a TBM's performance and are highly dependent on the machine's design parameters, geological conditions, slope and alignment of the tunnel, operational restrictions, managerial issues, and workers' experience of the project. Performance prediction of a TBM is one of the main factors for reducing the risk associated with the cost and schedule planning of construction projects [7,8]. Therefore, estimating TBM performance is critical for effective, economical, safe, and rapid tunnel construction [9,10].

Statistical models are less reliable in dealing with nonlinear and complex systems; extreme data values and outliners are often produced, making TBM performance prediction a big challenge [11,12]. However, flexible artificial intelligence (AI) techniques provide an opportunity to deal with highly complex and nonlinear engineering problems [8]. For predicting TBM performance parameters, such as PR, AR, etc., several AI techniques such as fuzzy logic [12,13], particle swarm optimization (PSO) [7,14], adaptive neuro–fuzzy inference systems (ANFISs) [12], imperialist competitive algorithms (ICAs) [8], gene expression programming [15], support vector machines (SVMs) [2], and artificial neural networks (ANNs) [7,8,16] have been used. Several optimizing algorithms have been integrated with SVMs to predict the TBM AR and energy consumption of the cutterhead drives in TBMs [4,17]. Similarly, ANNs have also been integrated with other algorithms to enhance the predictive performance of the system [7,8,18,19]. However, extracting and analyzing the required information from prediction analysis often requires extra effort for the stakeholder to visualize and understand the actual position and progress of a construction project. Therefore, a computer-based support system for better visualization of a tunnel construction project is needed to maintain dynamic information, which can automatically update potential information for stakeholders.

In the architecture, engineering, and construction (AEC) field, digital twins (DTs) are considered a comprehensive solution for presenting the lifecycle process of physical urban infrastructures in a virtual space, combined with the Internet of Things (IoT), big data, AI, and Semantic Web [20]. DT has been widely used in smart medical systems [21,22], product assembly shop floors [23–25], human–machine collaboration [26], studies on speed loss caused by marine fouling [27], additive manufacturing systems [28], engine optimization, maintenance [29,30], and product lifecycle management [31], blast design optimization [32], and other fields. Digital twins have been implemented for the lifespan prediction of perfricated components in noise barrier tunnels [33]. Highway tunnel pavement performance has been predicted using digital twins and time series stacking [34]. A decision analysis framework based on digital twins has been proposed for the operation and maintenance (O&M) of tunnels [20,35]. However, the TBM tunneling industry has yet to adopt digital twins. Therefore, we propose a digital twin-driven framework for TBM tunneling to predict, visualize, and monitor TBM performance through machine learning. It utilizes the dispersed big data from the TBM tunneling industry and eliminates the dependency on expert experience and knowledge to understand the construction progress. This study provides a solution and reference for TBM construction data analysis and machine learning model training for performance prediction. Furthermore, it provides a platform for the visualization and monitoring of TBM performance in virtual environments.

The rest of this paper is structured as follows: A literature review on tunnel construction machine learning for performance prediction and digital twins' potential use in tunnel construction is provided in Section 2. The architecture of our digital twin-driven framework assisted by machine learning for performance monitoring and visualization is described in Section 3. The methodology underpinning the TBM digital twin's modeling, machine learning algorithms, and training of algorithms is presented in Section 4. A theoretical case study as well as the results and discussion are presented in Section 5. Conclusions are outlined in Section 6.

2. Literature Review

The data generated during a tunnel project's site investigation, design, and construction processes vary in type, format, scale, and availability throughout the project [1]. When tunnel twin data are produced, the fusion of heterogeneous data from many sources is required at the data level. A computer-based decision support system is required to keep dynamic information and automatically uncover potentially valuable data to assist tunnel O&M decisions [20]. Building information modelling (BIM) is a sophisticated and parameterized digital modeling technique that can support the lifecycle activities of infrastructures in the AEC industry [20]. BIM can connect to several IoT data sources, as it has developed into an open platform for managing and sharing information. To provide more comprehensive decision support for the O&M of urban infrastructure, digital twins dynamically simulate physical entities' attributes, operating states, and evolution laws in the real environment. They also digitize expert experience that was previously impossible to preserve [36]. Thus, a well-defined "physical–data–virtual" framework and data mining techniques based on BIM and IoT can be established for higher interoperability, automation, and intelligence in delivering smarter construction services [37].

2.1. The Origin of the Digital Twin

DTs are regarded as an all-encompassing solution. In the discipline of AEC, researchers combine digital twins with the IoT, big data, AI, modeling analysis, and Semantic Web to simulate the lifecycle of physical urban infrastructures in a virtual environment [20]. Under the integration of physical products, virtual products, and pertinent connection data, digital twins generally refer to a mirrored digital depiction of the actual manufacturing process that can mimic all features of physical operations [38]. Michael Grieves first introduced the idea of digital twins in 2003 [39]. Although the concept of a "digital twin" was first presented in 2003, it is becoming more common in the current Industrial Revolution 4.0 [40–42]. More precisely, the National Aeronautics and Space Administration (NASA) study to continually simulate, anticipate, and analyze a spacecraft's state, seeking to reduce the degradation and failure in the vehicle, was particularly responsible for the resurgence of interest in digital twins [43]. Since then, experts have increasingly begun to pay attention to digital twins. The research company Gartner even listed the concept as one of the top ten most promising technological trends for the upcoming ten years in its 2018 forecast [44]. However, it has been observed that a constantly updated BIM data model lacks data manipulation capabilities to evaluate, assess, and forecast the current status of assets [38]. Besides the capabilities of BIM and DT models, AI provides the opportunity to assist stakeholders during complex decision-making by manipulating the data inflow to evaluate and predict the real-time status of assets, processes, and systems [45].

2.2. Machine Learning and Performance Prediction Applications

Artificial intelligence-based algorithms are used to explore the relationship between TBM performance and TBM performance parameters. An adaptive neuro–fuzzy inference system (ANFIS) that is more accurate than statistical models in predicting PR [9] has been developed. A new approach, group modeling of data handling (GMDH), was introduced to predict TBM PR accurately [46]. A gene expression programming equation was developed to estimate TBM PR accurately [15]. Several new optimization methods, namely hybrid harmony search, differential evolution, and the grey wolf optimizer, have been introduced to estimate TBM PR [14]. In another study, a hybrid SVM technique was used to estimate the energy consumption of cutterhead drives in shield tunneling [17]. Many ANN-based models have been developed for solving geotechnical engineering problems [46,47]. Hybridized models have been developed to predict TBM penetration and advance rates [7,8]. Researchers have performed various studies, developed modeling techniques, and identified input parameters to predict TBM performance in hard rock. However, AI-based models alone lack the understanding and visualization capability of tunnel progress and performance. With the assistance of DTs, AI can perform better in the overall performance prediction of TBMs.

2.3. The Digital Twin Applications in the Product Lifecycle

The digital twin system has been widely used in product design and production because it can help with understanding customer demands quickly, identifying or even foreseeing model weaknesses early, controlling production processes to quickly adapt to changing environmental conditions, and providing useful recommendations to optimize plant operation and maintenance before failure occurs [48,49]. In other words, the foundation for lining up the real world with virtual components is real-time data received from physical items. Automatic issue detection and performance evaluation allow for the formulation of optimized remedies that can be implemented in time to reap the rewards of increased dependability and efficiency [38]. A digital twin-based framework designed for the petrochemical industrial IoT and machine learning was presented to visualize the data exchange loop between the physical and visual model and production control optimization [48]. A digital twin-driven optimization method was presented to simulate a two-stroke heavy fuel engine, optimize its controlling parameters, and monitor its performance and manufacturing in real-time in a virtual environment [30,50]. A digital twin-driven intelligent predictive maintenance model was proposed to evaluate and monitor the degradation process and predict the remaining life of an aeroengine by employing deep learning [29]. Deep cyber–physical integration of intelligent manufacturing is now being pursued to increase production management flexibility, adaptability, and predictability [38].

Using smart inspection and component life prediction through digital twins benefits construction tunnel projects by reducing construction costs and labor efforts [51]. In another study, the prefabricated components of a tunnel were investigated to ease component life prediction by estimating displacement based on fitted data [33]. For unique circumstances, such as a fire breaking out in a tunnel, a digital twin-based robot system has been proposed to identify and track the location and status of the fire [33]. A real-time monitoring application for a TBM's interaction with the surrounding rock was developed to monitor cutterhead vibrations and shield jamming warnings [52]. However, there is still hesitation to apply DTs in the TBM tunneling industry, and there is still a gap in the literature regarding the visualization, monitoring, and performance prediction of TBM tunneling in a single platform.

2.4. Problem Statement and Objective

Due to the dispersed nature of the data accumulated from TBM tunneling, processing the input data and predicting the results, status of a TBM, and projection completion scenarios are difficult to understand for stakeholders. Furthermore, the prediction of TBM performance is a nonlinear and complex problem due to various geotechnical conditions encountered along the tunnel alignment [53]. Machine learning techniques have recently been adopted to solve this problem using only historical data to predict TBM performance. Feedback from real-time operational parameters and geological conditions has not previously been considered, which would influence the performance prediction accuracy of TBMs. An integrated application of digital twins and machine learning has the potential to solve such a problem.

This study proposes a digital twin-driven framework for TBM performance prediction, visualization, and monitoring through machine learning. The main contributions of this work are as follows: (i) based on different machine parameters and geological factors, different AI-based algorithms are used to predict the performance of TBMs using historical and real-time operational data; (ii) a virtual TBM model is constructed to provide an opportunity for end users to visualize, understand, and monitor each change in the state of TBMs; (iii) real-time information flow between physical and digital construction is established to ease schedule estimation; (iv) a framework is proposed to simultaneously predict, visualize, and monitor the performance of TBMs through digital twins and artificial intelligence.

3. Digital Twin Framework for TBM Performance Visualization and Monitoring

3.1. TBM Digital Twin

A digital twin can be adopted for high-quality transformation and digitalization in different TBM tunneling operations. It enables the system to virtually construct a digital replica of the actual system, stimulate and demonstrate the behavior of the system in a real environment while visualizing in a virtual environment, enhance the capabilities of the system by the addition of data fusion analysis, iterative decision-making, virtual–real interactive feedback, and machine learning techniques for the optimization and visualization of system performance. Considering the complex, heterogeneous, and dispersed operations of tunneling operations, data associated with TBM tunneling approaches, multiple data file sources, and the lack of visual capabilities of all this information at once are quite difficult for stakeholders to make sense of it. A TBM DT was constructed for the performance monitoring of TBM tunneling operations, as shown in Figure 1. The DT model consists of a physical and virtual part, the connection between the physical and virtual parts, and the data transfer between the models. The physical part is the TBM and data associated with its design and performance. The connection represents the directional flow of the data and describes the linkage between the different components of the TBM's digital twin-based platform. DT models provide a platform for visualizing each component of the TBM, visualizing its performance, and monitoring its functioning.

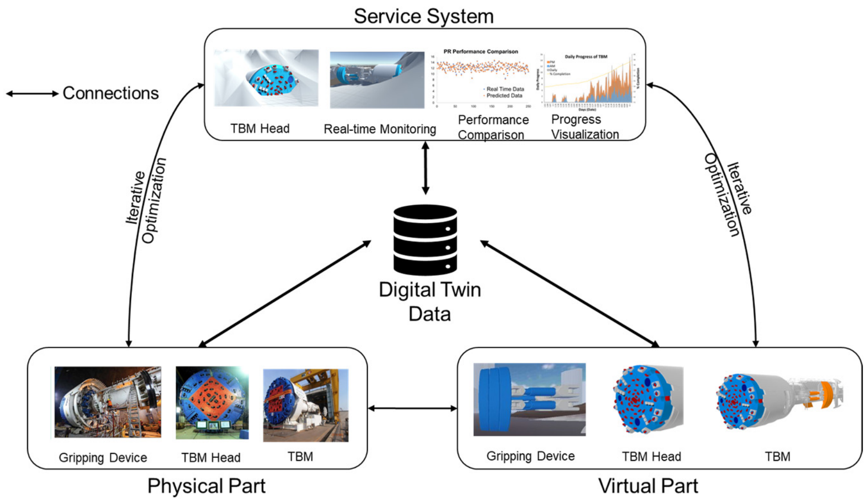

Figure 1. Digital twin structure of TBM performance monitoring.

3.2. TBM Digital Twin Structure

We developed a digital twin to create a digital replica of a physical entity in a virtual environment to simulate, predict, verify, and control lifecycle operation using historical data, real-time data, and machine learning algorithms. Digital twins, with the assistance of machine learning, provide a single platform to visualize and monitor performance, changes in TBM progress, and predictive results for stakeholders. Furthermore, the performance of the TBM is predicted through machine learning algorithms and linked with the database in the physical part. A schematic overview of the digital twin-driven framework for the visualization of TBM performance is shown in Figure 2. The main components of the five-dimensional digital twin framework are the physical part, virtual part, service, data, and connections.

3.2.1. Digital Twin Physical Entities

The physical world is the foundation of the digital twin paradigm. The physical entity of a digital twin could be any device or product, physical system, process, or organization. The TBM, the physical part of this study, is a complex machine with several different assemblies dedicated to specified tasks. The right type of TBM, assemblies, operating parameters, and, most importantly, proper linking of these assemblies is crucial for the desired performance of a TBM. The right selection of TBM type is important; still, the design of the TBM head in any geological condition is the most critical factor for the successful working of a TBM during tunnel construction projects. Cutterhead design includes information on cutter types, the spacing of selected cutter types for the given geological conditions along the tunnel, the profile and shape of the cutterhead, the balance of the cutterhead, the design and position of muck buckets, access to face, and cutting clearance for the cutters and body of a TBM. Due to the high dependency on the geological condition of a construction site, a slight change in cutterhead design may affect a project's cost and scheduling, TBM cutterhead stability, excavation performance, and the TBM's moveability.

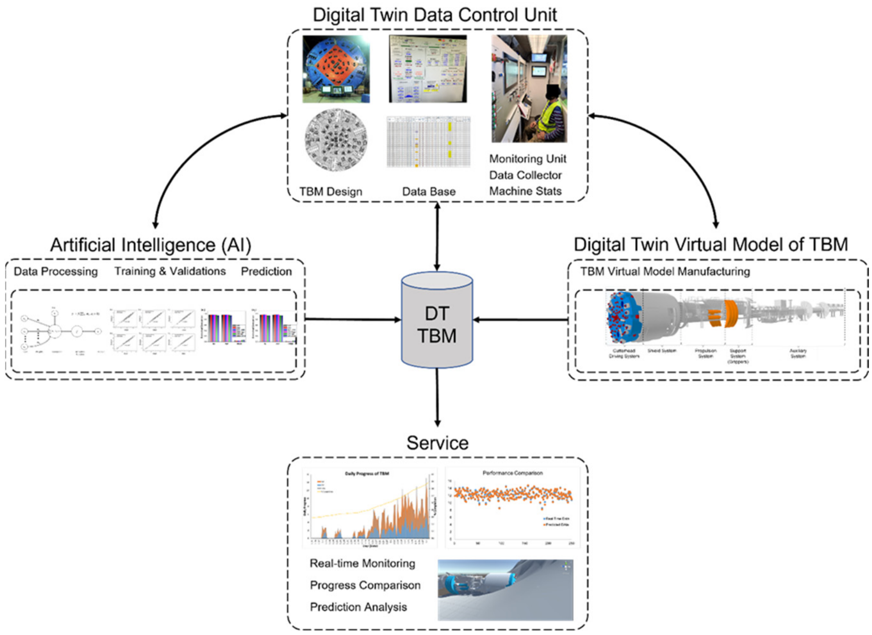

Figure 2. TBM digital twin structure.

3.2.2. Digital Twin Virtual Models

Virtual models are digital replicas of physical models to produce a real-world scenario in a virtual environment. Virtual models for digital twins are digital replicas of physical entities that reproduce physical geometry, properties, and rules. Detailed TBM design drawings were developed using computer-aided design (CAD), providing detailed geometric parameters for construction. Based on the design specification, a digital TBM replica was developed in a virtual environment. The three-dimensional digital replica model geometrically describes the physical entity in terms of size, shape, functions, and activities. Virtual TBM models facilitate the enhancement in visual monitoring capabilities regarding TBM performance.

3.2.3. Digital Twin Data

Digital twin data are considered the most important part of any digital twin model. The multidimensional, diverse, and heterogeneous nature of digital twin data varies for the physical and virtual models depending on the functions and activities to be undertaken. The TBM, also known as a moving factory, includes the cutterhead design, main drive, electric motor, conveyor belts, hydraulic propulsion and support system, etc. Detailed design drawings of the TBM, its components, and supporting assemblies were developed using CAD, providing detailed geometric parameters. TBM tunneling construction data and operating parameters were also recorded. Furthermore, feedback data from the simulation results, predictions from the machine learning algorithms, optimization, and design update results are also important for successfully implementing the digital twin data concept in a TBM tunneling project.

3.2.4. Digital Twin Services

Considering the complex and diverse data associated with the TBM tunneling method, it is difficult to visualize and understand the performance of a TBM system by following the traditional spreadsheet, 2D CAD drawing, or numerical data approaches. Service is an essential component of the digital twin paradigm. It provides the user with insight into digital twin application services such as monitoring, simulation, verification, and optimization results in a more presentable manner. The digital twin service system of this study is responsible for the visual representation of TBM performance, daily progress, and performance prediction results derived from the machine learning models.

3.2.5. Digital Twin Connections

Connections between physical entities, virtual models, services, and data enable information exchange. The digital twin's connections are responsible for describing the flow of digital twin data between different models. These connections are usually established between physical entities and virtual models, physical entities and data, virtual models and data, physical entities and services, virtual models and services, and services and data. Digital twin models are connected dynamically with their counterparts to enable the system to respond according to the requirement of making a digital replica of a physical entity or system. Furthermore, they also provide the same role while integrating with other applications, such as machine learning algorithms, for receiving prediction results.

3.3. Digital Twin-Based Framework for TBM Performance Visualization and Monitoring Method

Based on the proposed digital twin modules, we designed a digital twin-based framework for TBM performance visualization and monitoring. A digital twin and a machine learning-based six-step process was adopted to visualize and monitor TBM performance, as illustrated in Figure 3.

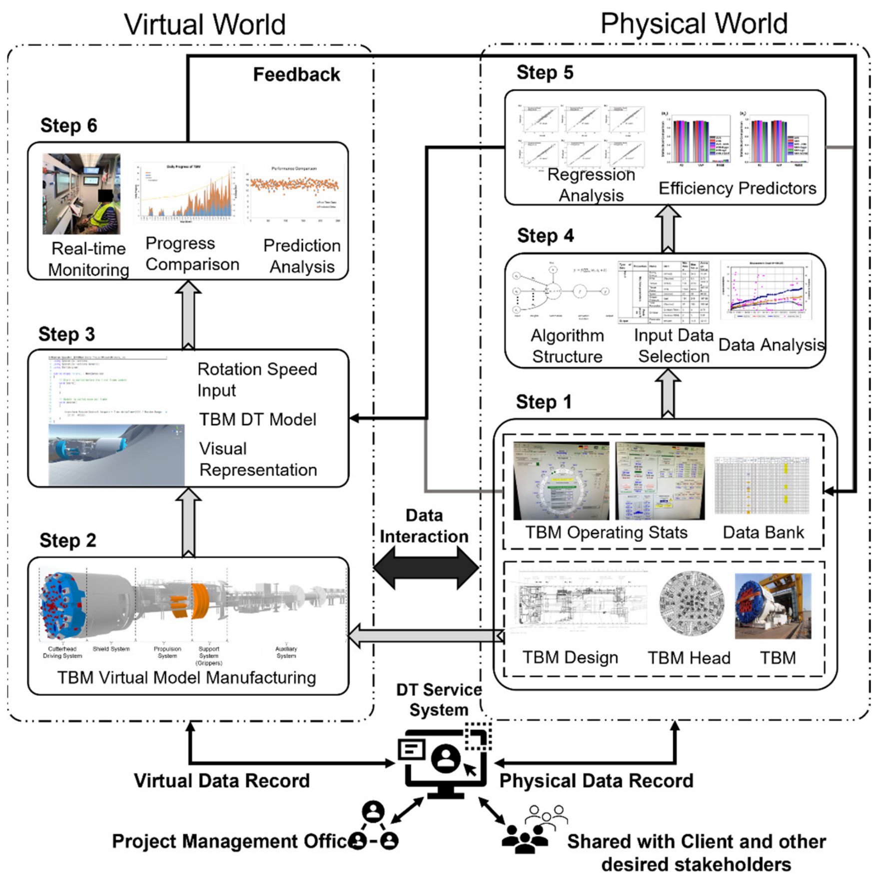

Figure 3. Digital Twin-based framework for TBM performance visualization and monitoring method.

Step 1: Understanding the geological conditions, rock mass mechanical parameters, TBM design parameters, and TBM operating parameters are key to successfully designing the right type of TBM, accurately executing TBM functions, and meeting the desired cost and scheduling of a tunnel construction project. A detailed feasibility study is used to determine rock mass mechanical parameters and geological settings. A pilot tunnel is constructed in the project area to have a detailed overview of the geological conditions, stresses, and deformation. A TBM is a moving factory that produces a tunnel of a desired geometric design. The tunnel's geometric design and parameters drive the TBM's geometric design. For that purpose, detailed design drawings of the TBM are developed using CAD, providing detailed TBM design parameters for tunnel construction. Furthermore, the operating parameters of the TBM are collected to store in the database for data visualization and provide an input feed for the machine learning algorithm for prediction and analysis.

Step 2: Based on the data from the previous step, TBM manufacturing is carried out in a virtual environment. The different TBM parts, such as the cutterhead, gripper, and main drive, are developed and assembled into a single TBM unit. Therefore, the design and assembly of each part of the TBM, especially the cutterhead, are implemented very precisely and accurately. Iterative design checks are conducted to obtain optimal geometric and motion parameters, and the digital mirror of the TBM is developed as per the requirement of the TBM digital twin for performance visualization and monitoring. Any change or error in the TBM design or assembly process incurs additional costs and causes a long delay in the project's schedule.

Step 3: After the virtual design, the TBM's geometry, parts, material, parts motion, and assembly process in a virtual environment are completed in step 2. The operating parameters of the TBM's operation during actual construction are selected from step 1 to feed the digital replica of the TBM constructed in the virtual environment. The digital replica is linked with step 1 to stimulate the operational condition in this step. The rotational speed of the cutterhead and other operational parameters of the TBM are controlled in the virtual environment. It is a virtual TBM product replicating and consistent with a TBM in a real-world environment in real time.

Step 4: This step mainly deals with data science and machine learning. The desired machine learning algorithms are selected, and their architecture and hyperparameters are optimized for the best possible analysis results. The operating parameters of the TBM and the geological factors of the ground are selected carefully, and the dataset is presented in a format suitable for the desired machine learning algorithm. The cutter rotation speed, torque, and thrust are the key parameters for the qualified TBM excavating the ground of interest. In the proposed framework, steps 3 and 4 execute their tasks simultaneously with steps 2 and 3.

Step 5: This step involves training and testing the machine learning algorithm to visualize the prediction performance of the TBM's operating parameters in real time. The model performance predictor results, root mean square error (RMSE), coefficient of determination (R^2), and variance accounted for (VAF) are estimated. The prediction analysis is evaluated and compared to find the best algorithm. Real-time feedback to the database in step 1 and the TBM digital twin are established simultaneously for design evaluation and performance visualization.

Step 6: Real-time monitoring is carried out to visualize and understand the TBM's performance and progress in a virtual environment close to the practical case, and feedback based on the TBM's performance is sent back to the database in step 1 to update the data. It helps the manufacturer quickly understand the required TBM design, operating conditions, tunneling operation progress, and assembly decisions for adequate project scheduling and financing.

4. Digital Twin Modelling Methodology

4.1. Data Collection

Data are produced from the tunneling survey sheets, geophysical surveys, borehole drilling and logging, in situ and laboratory testing, and geological surveys. They also include the operating parameters of the TBM, equipment electrical and mechanical status, and status of the TBM components. The type and format of the data differ in nature and size, including images, spreadsheets, and text documents. The TBM data are continuously generated. Furthermore, they include the geological data during the excavation. The data measured using sensors and other equipment are recorded and observed in the TBM cabin. The TBM continuously receives the data and stores them in a specific order and continuous number. The program for data acquisition automatically starts after each new start of the PC and loads all required program components into storage. Afterwards, it determines the instantaneous operating state of the machine. Termination of the program is controlled through the control cabin with the switch cabinets of the control panels. All the main controls of the TBM and components are connected using cables, tested, and started up; furthermore, a connection to the PLC of the TBM is established. In the switch cabinets, the programmable control system, data acquisition, as well as measurements and data cards are integrated. The programmable logic controller (PLC) is an S7 type, manufactured by Siemens. The control panels house all the necessary operating and display units installed. The operator panel desk is IP48-rated. A visual display shows all operational parameters of the TBM and its systems. Local control panels are located adjacent to moveable parts of the machine, and all are fitted with emergency stops and key switches where appropriate, for local isolation during maintenance. A Siemens S7 PLC is at the heart of the control system and is provided to control the main functions of the TBM; it is installed in the operator control desk with remote interface units installed in the distribution panels. The PLC system is interfaced with an industrial computer within the operator control desk. A graphical representation of data transmission from the TBM to the monitoring device is shown in Figure 4. A controller area network (CAN) bus uses a two-wire bus architecture to transmit and receive data messages between devices. It is a message-based protocol, which makes the CAN bus highly scalable and efficient in handling large amounts of data in real time. The software is protected from unauthorized access. The PLC system has a powered battery for power loss, a fail-safe system, interlocked circuits, and critical safety circuits hardwired separately to the PLC.

It is necessary to investigate the effect of all the influential parameters and choose the key parameters important for a better AI-based predictive model of TBM performance. All influential parameters can be categorized into three related rock properties, machine characteristics, and tunnel geometry [12]. Based on the literature, compressive strength is the most influential parameter affecting TBM performance [53]. In another study, RMR, RQD, and UCS were considered the most influential parameters affecting TBM performance [16]. Similarly, an empirical TBM performance model has also been presented to predict PR using the Q and QTBM values [54]. Other researchers also have included the joint conditions, BTS, and weathering degree of the rock mass to predict TBM performance [16,55]. Besides these rock-based parameters, thrust force (TF) is the most important machine parameter for predicting TBM performance [2,12]. The maximum torque, maximum power, maximum revolutions per minute (RPM), and the function of thrust force are important parameters for TBM performance. Since the geometry parameters usually remain constant, these parameters are not considered in this study. A degree of multicollinearity among different parameters such as uniaxial compressive strength (UCS), rock quality designation (RQD), rock mass rating (RMR), and Q-value may affect TBM performance. Although each of these parameters provides key information regarding the rock characteristic but avoids the multicollinearity and computational complexity of the predictive model, the Q-values are selected for the input parameter. In previous studies, several machine parameters have been selected for the TBM performance prediction model, such as thrust, torque, boring energy, gripper pressure, and shield pressure. The parameters we selected for predicting TBM performance are boring energy, RPM, torque, TF, speed, gripper pressure, total revolutions, Q-value predicted by the site engineer, and TBM. The data collection of TBM input parameters from the Neelum Jhelum Hydroelectric Power (NJHEP) project site was carried out carefully, as shown in Figure 5. Figure 5a shows the inside of the TBM control cabin displaying and recording the real-time machine parameters, Figure 5b shows the real-time

TBM screen displaying machine operating data, and Figure 5c,d shows the TBM thrust cylinder, rotation motor, and boring data during operation, respectively.

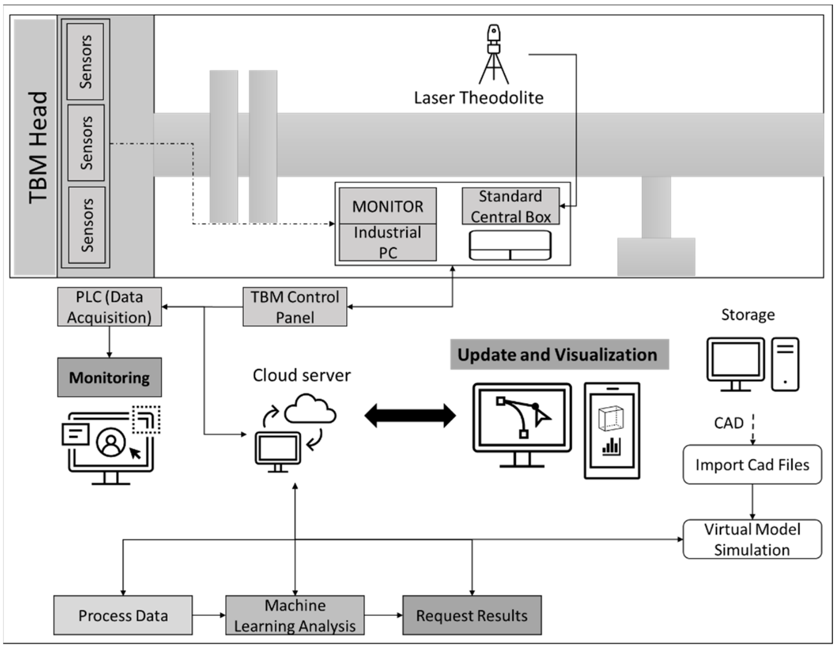

Figure 4. Data transmission from TBM to monitoring.

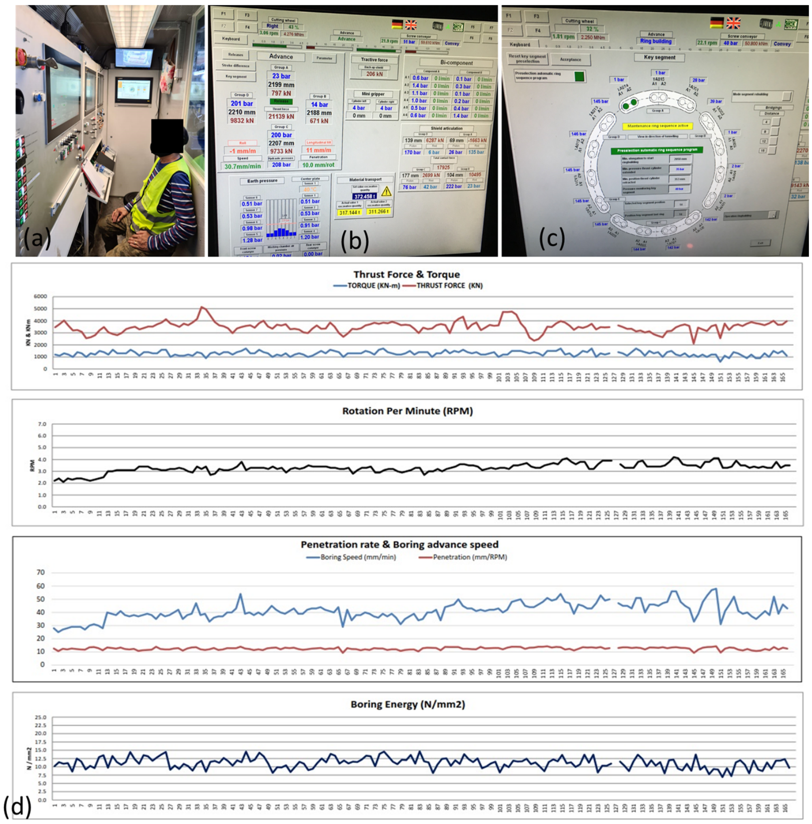

Figure 5. Data collection and TBM input parameter of Neelum Jhelum project: (**a**) TBM control cabin showing and recording the real-time machine parameters; (**b**) snapshot of real-time TBM screen showing working machine data; (**c**) thrust cylinder and rotation motor data; (**d**) boring data of TBM during operation.

4.2. Input Parameters

4.2.1. Machine Performance Parameters

The influence of effective machine parameters on the TBM's performance prediction of is vital. PR is defined as a ratio of distance bored to the boring time during tunnel construction (Equation (1)), where PR is the average penetration rate at which the cutterhead bores rock per hour or minute, expressed in m/h or mm/min. Thus, the operating time is used to calculate the PR to measure the cutterhead advance per unit of boring time.

$$PR_{TBM} = {}^{\text{Distance bored}}/_{\text{Boring time}} \tag{1}$$

$$P_{rev} = {}^{1000 \times \text{PR}}/_{60 \times \text{RPM}} \tag{2}$$

P_{rev} is penetration per revolution of cutterhead (Equation (2)), and RPM is the rate of cutterhead revolutions per minute, expressed in mm/rev and rev/min, respectively. AR is the ratio of both mined and supported actual distance to the total time. The key parameters for TBM performance prediction used in this study are boring energy, RPM, torque, TF, speed, gripper pressure, and total revolutions.

4.2.2. Geological Rock Parameters

The Q-value describes the stability of the underground opening. The range of the Q-value is a numeric value on a logarithmic scale and lies between 0.001 and 1000, denoting very low to exceptionally good rock quality [56]. RQD, joint number (Jn), joint roughness (Jr), joint alteration (Ja), joint water reduction factor (Jw), and SRF are the deciding factors of the Q-value prediction of the rock. Based on these factors, the Q-value is predicted by Equation (3) [57].

$$Q = \frac{RQD_0}{J_n} \times \frac{J_r}{J_a} \times \frac{J_w}{SRF} \tag{3}$$

The determination of the Q-value is made possible using geological mapping in the underground structure, on the surface, or using core logging. The true Q-value determined in the excavation site is usually more accurate. However, the rock parameters are incorporated with the already available Q-value to predict TBM's performance accurately in Equation (4) [54].

$$Q_{TBM} = \frac{RQD_0}{J_n} \times \frac{J_r}{J_a} \times \frac{J_w}{SRF} \times \frac{SIGMA}{{}^{F^{10}}/_{20^9}} \times \frac{20}{CLI} \times \frac{q}{20} \times \frac{\sigma_\theta}{5} \tag{4}$$

where F = average cutter load through the same zone, SIGMA = rock mass strength estimate (MPa) in the same zone, CLI = cutter life index, q = quartz content in percentage terms, and σ_θ = induced biaxial stress on tunnel face. In this study, two Q-values are used to predict TBM performance. The site engineer determines one Q-value using geological mapping, while the TBM predicts the other. Using these two values for predicting PR compensates for most of the geological parameters influencing the TBM's performance.

4.2.3. Data Normalization

The normalization process of the TBM data is needed to obtain more effective general input data for AI algorithms and reduce the job site-dependent characteristics. AI algorithms (ANNs) are usually sensitive to feature scaling, so it is highly recommended to scale the data. For this purpose, the dataset is scaled to each attribute, either input vector X to [0, 1] or [−1, +1] or standardized to have a mean of 0 and variance of 1. Applying the same scaling to the complete dataset (training and test dataset) is recommended for meaningful results. The Standard Scaler was used to standardize this study, providing the zero (0) mean. An overview of the input and output dataset is shown in Table 1.

Table 1. Overview of selected TBM data.

Type of Data	Properties	Name	Unit	Min Value	Max Value	Average Value
Input	Machine parameters	Boring energy	(N/mm^2)	0.9	24.9	11.97
		RPM	(Rev/min)	2.1	5.4	3.70
		Torque	(KN-m)	100	2700	1336.34
		Thrust force	(KN)	1502	8239	3871.56
		Speed	(mm/min)	25	58	44.98
		Gripper pressure	(bar)	184	249	197.65
		Total revolution	(Rev/min)	61	188	106.94
	Rock properties	Q-value	Q-value (theo)	3	4	3.75
			Q-value (TBM)	2	5	3.91
Output		Penetration	mm/rpm	8	14.8	12.19
		Advance rate	m/h	1.26	3.69	2.73

4.3. TBM Twin Modelling

The TBM design includes the data from 2D and 3D design charts, design dimensions of electric and mechanical devices, CAD, and sensor-based data of the physical TBM for the validation purpose of the TBM model. A detailed TBM design is shown in Figure 6. The physical model of the TBM is a general hard rock open gripper TBM with a tunnel diameter of 8.5 m selected for the description of the framework. The TBM has hard rock drum-type cutterheads with single- and double-disc cutters. The cutter shield has a design load of 300 KN/m^2 with a boring stroke length of 1800 mm, having four propelling cylinders. Data are the core elements in the digital twin paradigm. The goal is to obtain precise and accurate information to better develop the digital twin-based TBM model and record the data during tunnel construction. The TBM design includes the data from 2D and 3D design charts, design dimensions of electric and mechanical devices, CAD, and sensor-based data of the physical TBM.

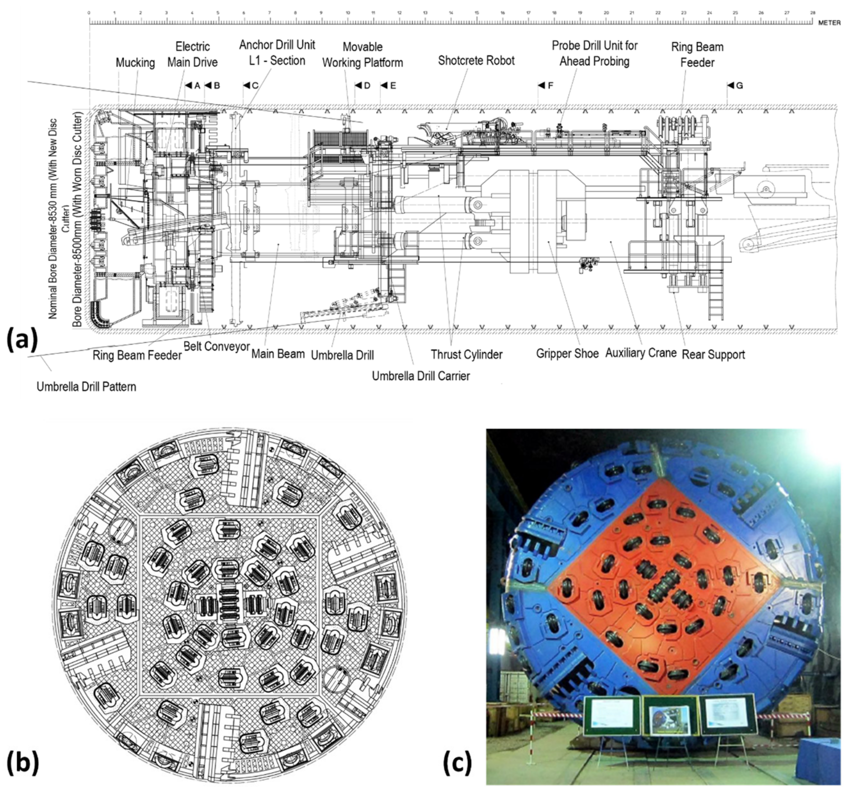

Figure 6. Details of TBM 696: (**a**) tunnelling system, components, and their arrangement; (**b**) cutterhead with the arrangement of components; (**c**) TBM 696 photograph.

The TBM structure created for the program in a gaming environment and Python consists of several steps. First of all, the required libraries are requested, and then, using the required libraries, all the variables involving the process are declared, such as working time, simulation time, etc. These variables are used in all the processes performed by the person operating the machine. The process methods are defined for every process to obtain the times for the simulation, and the machine environment is created, in which all the states and variables are set to their initial positions: working and not working. Furthermore, the methods in the machine environment are defined for required modification in the machine operation. Finally, the results, such as the rotation speed, and cycle time are shown. The TBM digital twin model is linked with the cloud database and depicts the actual progress in the virtual environment. A brief overview of the visual demonstration is shown in Figure 7.

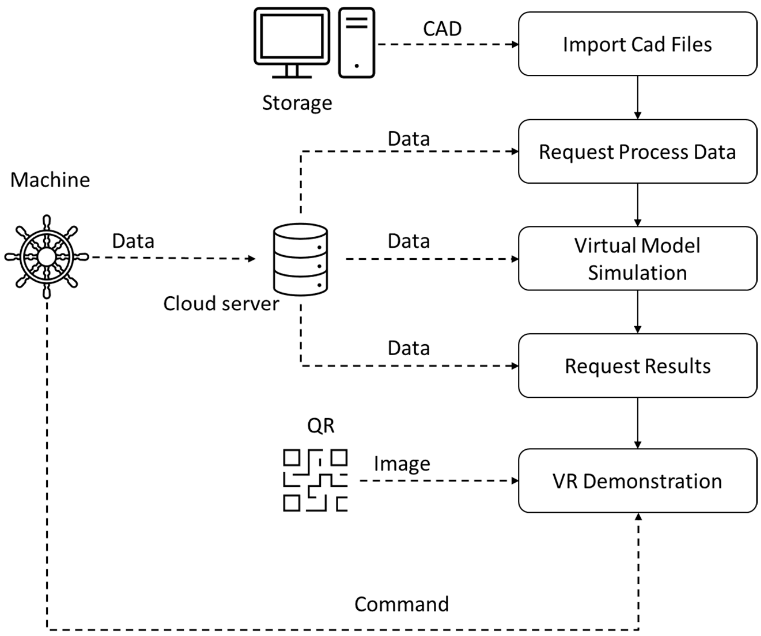

Figure 7. Digital twin modeling for visual demonstration.

The most important part of the TBM is the cutterhead design, which is very sensitive to the performance of the TBM. The cutterhead is equipped with seventeen-inch disk cutters, including fifty single- and four double-disk cutters in the center. The rotation of the cutterhead causes the discs to roll in concentric tracks on the tunnel face. The cutterhead also contains six buckets. The excavated muck is collected by the buckets with the rotation of the cutterhead and deposited on the belt conveyor via the muck ring. The cutterhead is equipped with ten water injection nozzles to reduce the temperature and dust formation at the tunnel face. The main drive consists of the housing, main bearing, ring set, pinions with bearings and shaft, planetary gears, electric motors, safe sets, sealing systems (inner/outer), and brake. The main drive includes 12 fixed installed electric motors of 350 kW each. The grippers are usually used in the hard, stable rock condition to provide the thrust to the cutterhead for a 2–12.5 m excavation drive by exerting the shove forcing against the tunnel wall through a hydraulic gripper reaction system. The hydraulic gripper-designed TBM is designed for a boring stroke of approximately 1.8 m.

The TBM digital twin model was created based on physical TBM design information. In this research, using Unity3D (Unity, Vesrsion 2020.3.15f2), a gaming environment software package, a 3D TBM model was created, as shown in Figure 8. For the one-to-one real mapping of the virtual environment, and to ensure the physical dimensional relationship between the virtual and real world, a CAD model developed based on physical measurements was placed in the virtual environment. The model is a replica of the physical model. The digital twin of the virtual environment is a digital mapping of the TBM's status, progress, and performance during the lifecycle of the project. The developed virtual environment allows for interaction with TBM elements using Oculus Quest 2 hand controllers. It also allows for users to navigate in a virtual environment by responding with physical body and head movements. The main components of a TBM include the cutterhead (cutter driving system), shield system, propulsion system, support system (grippers), and auxiliary system to handle TBM requirements.

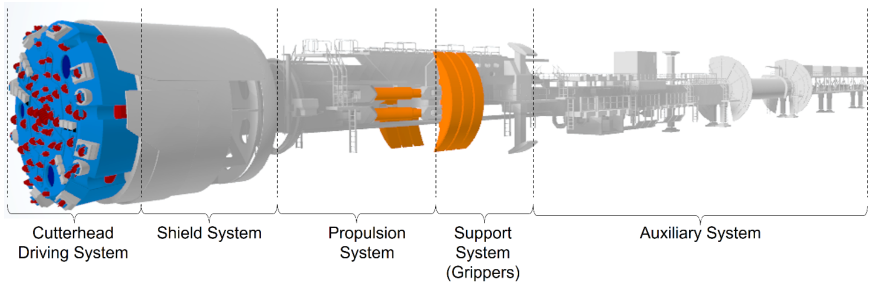

Figure 8. Three-dimensional virtual model of a TBM and its components.

4.4. Model Training Algorithm

4.4.1. Support Vector Regression (SVR)

SVM is a supervised learning method widely used in statistical classification and regression analysis. The SVM model provides the best separation hyperplane in feature space (FS) such that the negative and positive sample intervals in the training set are the largest. With the help of the SRM inductive principle, SVR improves its generalization ability using only a limited number of patterns [2]. Therefore, the SVM technique as a regression analyzer, SVR, is applied in this research to tackle regression estimation problems for TBM performance prediction.

The main objective of SVR is to find a function f(x) with the minimum deviation between the predicted and actual values. SVM makes all sample points approach the hyperplane by minimizing the total deviation between the sample points and the hyperplane [58]. The linear approximating function of SVR, $f(x_i)$, is represented in Equation (5).

$$f(x_i) = (w_i, x_i) + b \tag{5}$$

where w identifies the weight vector having a unit length at a right angle with the hyperplane and bias (b) corresponds to the threshold coefficient.

4.4.2. Artificial Neural Network (ANN)

ANN is a control AI technique that attempts to mathematically model and stimulate the computational relationships between data by replicating the reasoning operation of the human brain [7,59]. It allows one to handle nonlinear problems better than classical analysis methods [60]. An ANN usually consists of three layers: input layer(s), hidden layer(s), and output layer(s). The net input of each neuron is processed using an activation function (e.g., sigmoid or rectified linear unit (ReLU)). The weighted input signal is achieved by multiplying each hidden neuron's total net input (x_i) from the previous layer with an adaptive weight coefficient (w_i). The summation function of these weighted input signals, plus a small amount of bias (b), repeatedly for each neuron and layer provides the overall system output [7]. A systematic overview of the process is shown in Figure 9. Limited memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) is an optimization algorithm used in this study that belongs to the quasi-Newton method family. It approximates the BFGS algorithm using a limited amount of computer memory. It finds a (local) minimum of an objective function by using objective function values and the gradient of the objective function. The advantage of L-BFGS is that it only retains the most recent gradients, which is a much smaller storage requirement than the full Hessian estimate, as is required with BFGS. Unlike (full) BFGS, calculations required to estimate the Hessian in L-BFGS are accomplished without explicitly forming it.

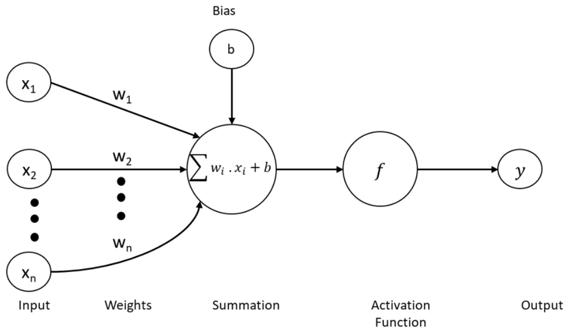

Figure 9. The basic architecture of an artificial neural network (ANN).

4.4.3. Integrated SVR-ANN Model

Researchers have used numerous alternatives to improve the performance and generalization capabilities of ANNs using other algorithms in engineering and science problems. All ANN-based predictive models tend to minimize cost function by adjusting weights and biases. We propose a hybrid SVR-ANN predictive model to predict tunnel performance. SVR is used here to extract the feature of input feature space by using an optimal hyperparameter value. The optimal values of the parameters selected based on the producing maximal accuracy were the most appropriate parameters. The optimal values were then used to extract the features using the SVR model. The extracted features are removed if the corresponding importance of the feature values is below the threshold parameter. The output of the SVR model is fed to the ANN model. Again, the ANN model's optimal hyperparameters were selected based on the producing maximal accuracy of the training function. The process of information flow of the SVR, ANN, and SVR-ANN integrated model is described in Figure 10.

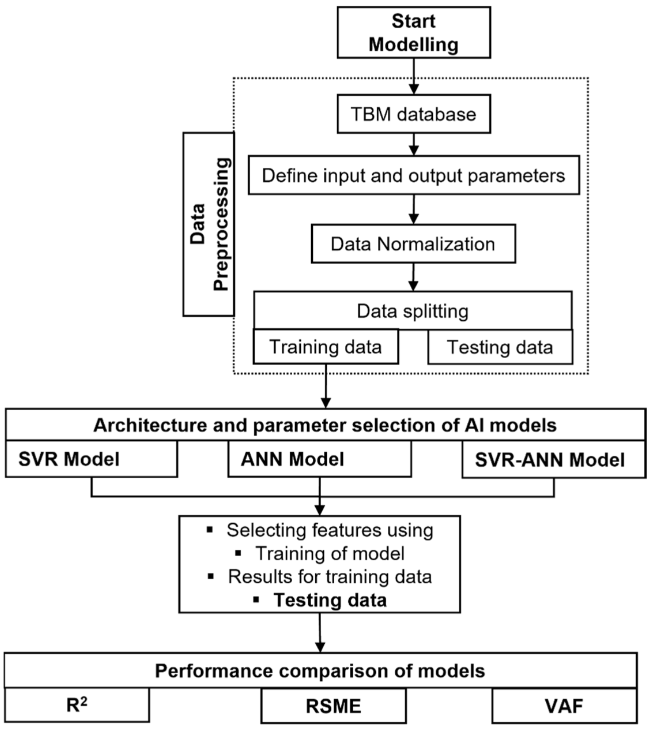

Figure 10. Flowchart of information flow of proposed models.

4.5. Model Training

A TBM is a complex system that undergoes operations such as muck transport, ventilation, and ground support, providing rotational stability under different geological conditions and mass rock properties. This research covers the mechanical characteristics of the machine and geological conditions incorporated in the SVR model to enhance predictive accuracy. This paper implemented three algorithms (i.e., SVM, ANN, and SVM-ANN) to explore better methods for predicting TBM PR. Then, these three intelligent models were constructed using the training set. In the above optimization process, different hyperparameter configurations and different model prediction performances were obtained. The SVR represents the complex function depending on the number of support vectors irrespective of the input space's dimensionality. However, the selection of control parameters is crucial in SVR compared to other AI algorithms.

Selecting optimal design parameters is a key step for the better performance of the SVR model. Generally, three parameters (the loss function parameter, the regularization term, and the Gaussian kernel parameter) should be selected appropriately for training an SVR model. The regularized constant (C) plays the role of minimizing the error term. The optimizer selects a smaller margin hyperplane to better predict results for a large value of C. The free parameter used in the model is C, and the "RBF" is the Gaussian kernel function used in this study. The parameter y denotes the variance in the Gaussian kernel, controlling the sensitivity of the kernel function. The MLP regressor iteratively trains the partial derivatives of the loss function at each time step to update the model parameters. A regularization term added to the loss function shrinks model parameters to prevent overfitting. The activation function, ReLU, and the solver, L-BFGS, used "for weight optimization" belong to the family of quasi-Newton methods used in this study. A total of nine (9) input layers, two (2) hidden layers, and one output layer are used in the architecture of this study. The number of neurons in the first and second layers is 150 and 100, respectively.

Parameter C is responsible for the smooth decision boundary and accurate classification of the training data points. For that purpose, the controlling parameter C (penalty parameter of the error term) for the SVR model varied from 0.1 to 100 to analyze the effect of change in prediction performance and the results were collected. Similarly, the controlling parameter alpha (α) of the ANN model varied from 0.1 to 100 to understand the variation in the prediction of results. Similarly, the α is a regularization parameter responsible for better fitting the prediction results by controlling the size of the weights. The comparative results of SVR and ANN at different C and α are presented in Figure 11. R^2 and RMSE values of these algorithms were compared at C and alpha values from 0.1 to 100. It is shown that both models show the best results at C = 10 and α = 10. Therefore, these values are taken as an optimum value to conclude the results of the predictive models.

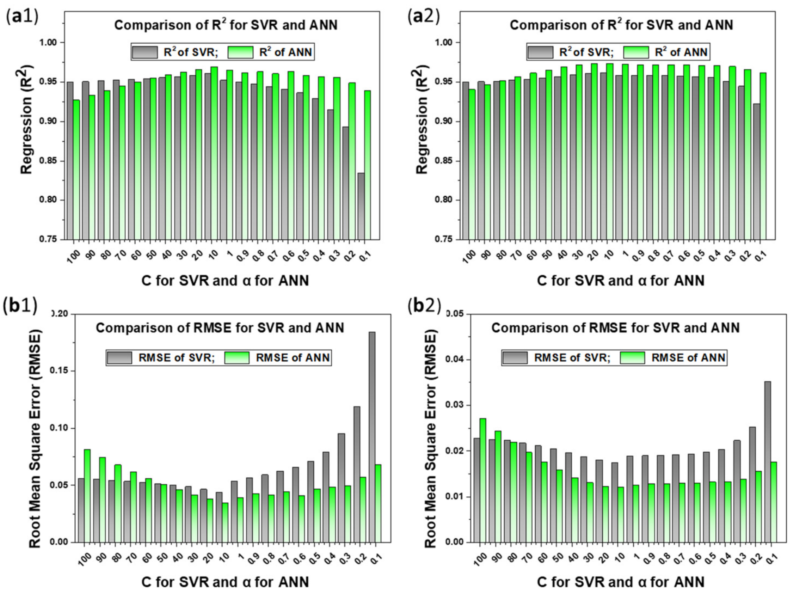

Figure 11. Optimization of SVR and ANN models at different C and α values: (**a1**,**b1**) comparative results of predictive models for PR; (**a2**,**b2**) comparative results of predictive models for AR.

4.6. Model Performance Predictors

At the last stage of model development, we ran evaluations to predict TBM penetration and advance rate. This section compares the models mentioned above to choose the most efficient one. The results obtained from these models were examined according to performance indices based on the statistical parameters, including RMSE, R^2, and VAF. The performance of the model was evaluated using RMSE, R^2, and the VAF in percent (%), using Equations (6)–(8).

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum\nolimits_{i=1}^{N}(y - y')^2} \tag{6}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y - y')^2}{\sum_{i=1}^{N}(y - \tilde{y})^2} \tag{7}$$

$$\mathrm{VAF} = \left[1 - \frac{\mathrm{var}(y - y')}{\mathrm{var}(y)}\right] \times 100\% \tag{8}$$

y and y' are the measured and predicted values of output, respectively, $\tilde{y}$ stands for the mean of the variable y, and N is the total number of datasets. Based on the created datasets, ANN models were developed to predict TBM PR, and their evaluations were performed based on R^2, RMSE, and VAF. To obtain a theoretically perfect network model, RMSE, R^2, and VAF should be 0, 1, and 100, respectively.

5. Theoretical Case Study, Results, and Discussion

5.1. Project Overview

TBM predictions help control the construction risk related to cost, operation, and decision under certain conditions. Machine and rock parameters should be incorporated into the model to predict the performance of TBMs close to reality. This research collected data from the NJHEP project, the second largest hydropower tunneling project globally (Figure 12). Two TBMs were used to excavate twin tunnels, each 10 km in length. TBMs were adopted to enhance the project's productivity, but due to the complex nature of geology, the TBMs faced serious problems [61]. The TBM tunnels are driven in the central portion of a zone bounded by two major Himalayan faults. The first is the Main Boundary Thrust Fault, and the second is the subsidiary reverse or Muzaffarabad Fault. The first fault extends through the mountain range's length following the Neelum River's course at the upstream start of the headrace tunnels at Nauseri, while the second fault runs close to the course of the Jhelum River at Thotha. A rupture along this fault in 2005 resulted in the Muzaffarabad earthquake, which caused over 75,000 fatalities.

The project is located at the foothills of the northwestern Himalayas (Murree Formation), where the infrastructure is inadequately developed. The Murree Formation comprises tight folding with repeated faulting and fracturing, indicating a high structural compression. Three detailed laboratory testing programs were conducted at the feasibility/detailed design, commencement of construction, and construction phase to determine rock parameters. Three main rock units of the Murree Formation, sandstone, siltstone, and mudstone, were found and classified during the detailed design stage. Sandstone is well cemented, generally strong, fine- to medium-grained, greenish-grey to grey, and moderately to closely jointed. Siltstone is strong to medium-strong, greyish brown to brown and reddish-brown, and closely jointed. Mudstone is weak to medium-strong, fine- to very fine-grained, reddish-brown, and closely jointed. Table 2 summarizes some of the intact rock mass parameters of the Murree Formation.

(a)

(b)

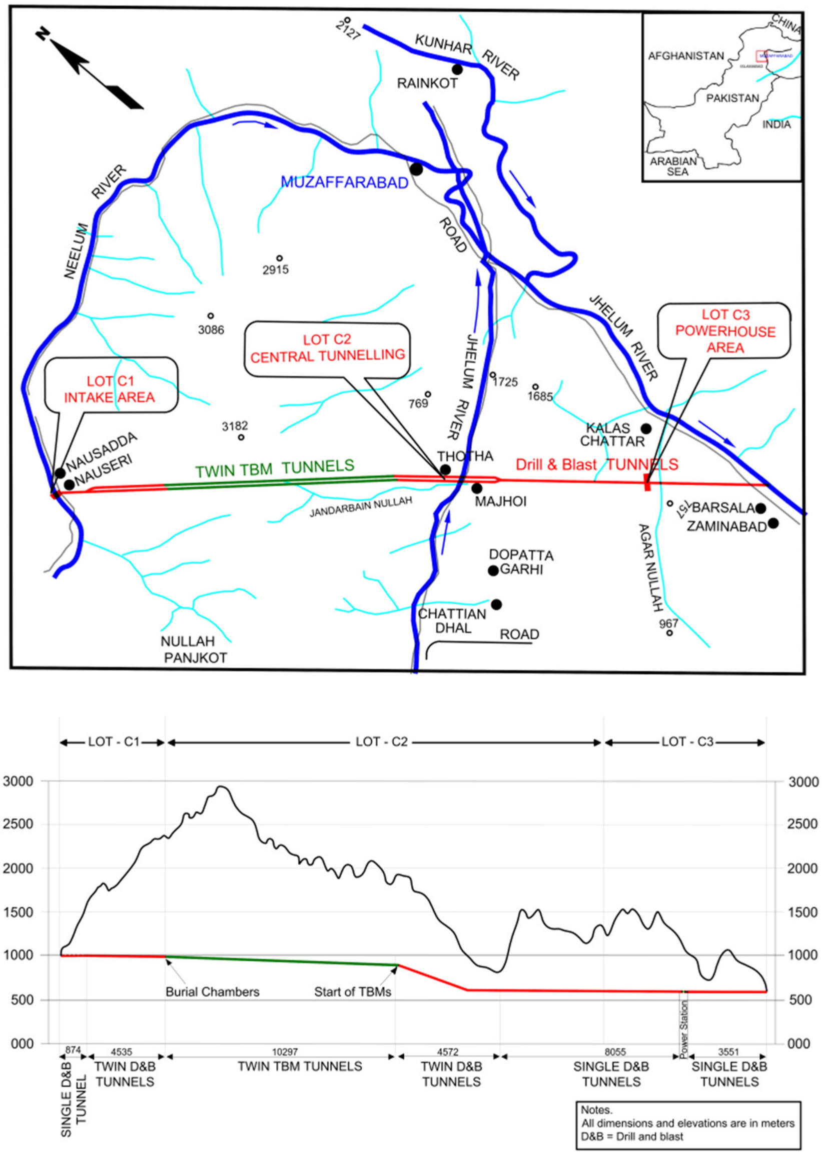

Figure 12. Neelum Jhelum project: (**a**) layout showing the alignment of TBM twin tunnels and other marked features; (**b**) terrain profile of the tunnel.

Table 2. Intact rock parameters of the main rock types.

Parameter	Sandstone	Siltstone	Mudstone
Color	Gray	Brown–reddish-brown	Reddish-brown
Weathering	Fresh–slightly	Fresh–Slightly	Fresh–slightly
Structure	Massive, blocky, locally irregular	Blocky, Tabular, Locally Irregular	Tabular, blocky, irregular
Grain size	Fine–medium	Very fine–Medium	Very fine
Bedding	Thick–massive	Thin	Very thin–yhin
Bulk density, kg/m^3	2730	2771	2722
Uniaxial compressive strength, MPa	86.0	56.5	33.0
Average rock quality classification	Good	Fair	Poor
Volumetric joint count (joints/m^3)	1–22	3–25	3–25
Number of joints sets	2 + random to 3 + random	3 to 3 + random	3 to 3 + random
Joint roughness, waviness	Rough, planar–undulating	Rough–smooth, planar	Smooth, planar
Joint aperture or thickness (mm)	<0.1–10	0.1–10	0.25–5
Joint filling	Clean, sandy particles, or hard calcite	Clean, sandy, or silty coatings	Silty or clayey coatings, occasionally soft clay

5.2. Result and Discussion

Based on the TBM operating parameters and geological parameters of the rock, a machine learning algorithm was applied to predict TBM performance in advance and help visualize the project progress status. The most important TBM parameters during tunnel excavation are advance speed (mm/min), rotational speed (rpm), advance pressure (bar), cutterhead torque (MNm), total advance force (kN), penetration (mm/rot), the pressure of crown–support–cylinder left and right (bar), and path of crown–roof–support–cylinder (mm), which are recorded at a specific interval during the excavation process of the tunnel. All this information is recorded to systematically understand the behavior of the TBM. Also, it includes the geological data, TBM's geographic location, daily progress, and remaining tunnel excavation length. The daily progress of the TBM over a month is shown in Figure 13.

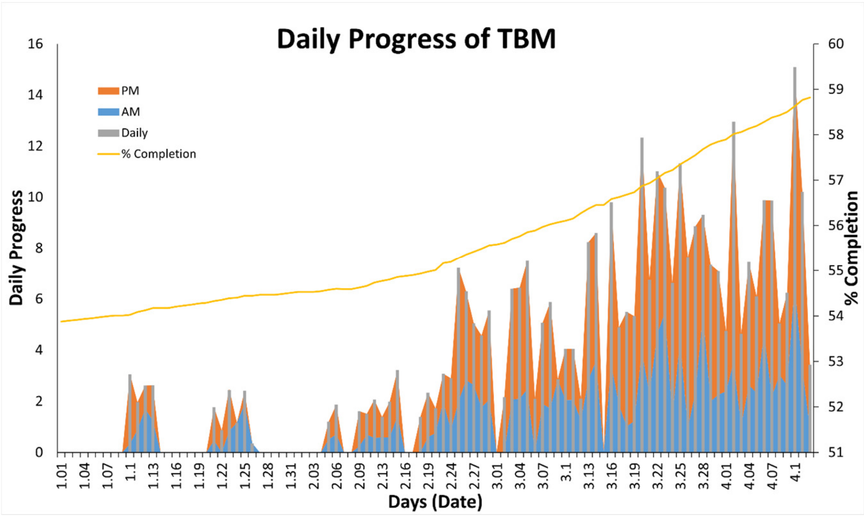

Figure 13. TBM daily progress (in meters).

A digital twin-based framework for TBM performance prediction, visualization, and monitoring was designed with machine learning. Based on TBM design and operational information, TBM manufacture was carried out in a virtual environment. It is a virtual TBM product that replicates a real TBM and its behavior in a virtual environment to help the manufacturer quickly understand TBM design and decision-making for project scheduling. For that purpose, a digital twin and a machine learning-based six-step process were adopted to visualize and monitor TBM performance on virtual platforms. The operating parameters were then transferred to the virtual model to visualize the operating condition. The operating parameters of the TBM and the geological factors of the ground were selected carefully to prepare the dataset in a suitable format for the desired machine learning algorithm. A machine leaning algorithm was developed to feed the virtual model with different RPM values and the revolution of the machine was demonstrated in a virtual gaming environment. Figure 14 shows the TBM virtual model rotating at the given RPMs. Real-time monitoring was carried out to visualize and understand TBM performance and progress in virtual environments close to the real case. Feedback based on TBM performance was sent back to the database to update the data. The TBM's historical and current operating parameters were collected for visualization, performance prediction, and progress comparison. This provides an opportunity for the end user to understand and monitor the state of the TBM and real-time information about the system, subsystem, and key components by forming a closed loop of a physical TBM, virtual TBM model, digital twin data platform, and the virtual environment to visualize each change in the state of the TBM.

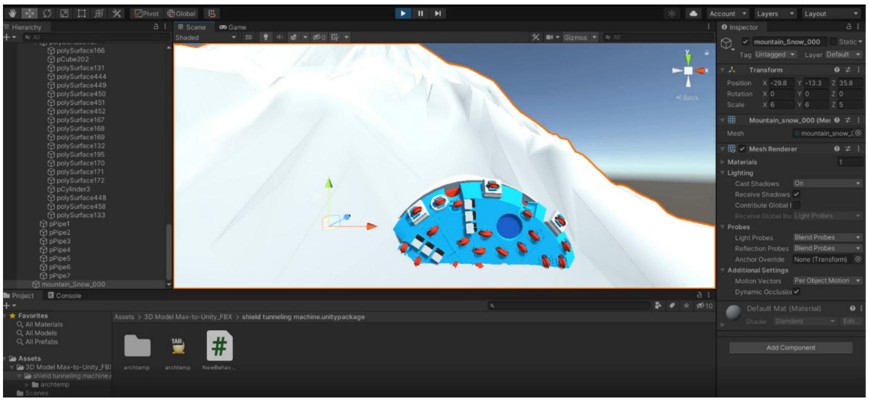

Figure 14. Practical demonstration of TBM rotation.

TBMs have many advantages, they still encounter many problems due to complex geological conditions. AI-based algorithms such as SVR can test the feasibility and applicability of a TBM before project commencement. Using the SVR model, TBM penetration rate predictions can be made to an adequate level, which helps control cost, as the interaction between machine and rock mass can be accommodated. The efficiency of the SVR model is estimated by comparing the measured and predicted values. The SVR model shows an accurate PR prediction compared to the measured ones. The predictive performance indices RMSE, R^2, and VAF are computed to evaluate the efficacy of the model. The performance indices RMSE, R^2, and VAF are predicted quite accurately, as shown in Figure 15(a1,a2) and Table 3. The findings of the SVR predictions show high conformity between predicted and actual SVR values of TBM PR. Similar results were reported in [2] using a limited dataset. Recently, authors of [18] presented promising results by using rock strength properties as input parameters.

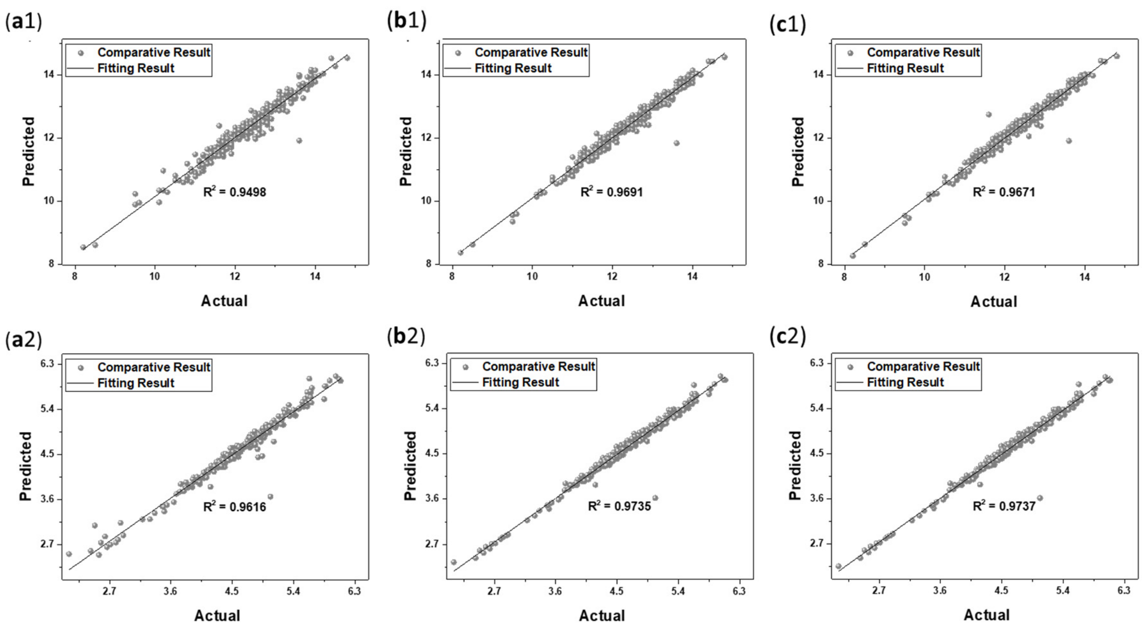

Figure 15. Regression analysis based on the actual and predicted TBM values; (**a1**,**b1**,**c1**) represents the penetration rate while (**a2**,**b2**,**c2**) represents the advanced rate for SVR, ANN, and SVR-ANN models, respectively.

PR prediction using the ANN model was carried out using L-BFGS as a solver function. The controlling parameter α for the ANN model varied from 0.1 to 100 to find the best prediction results, as shown in Figure 11. The predictive performance indices RMSE, R^2, and

VAF of ANN models showed improvement. The correlation between predicted and actual TBM PR values for the ANN is plotted in Figure 15(b1,b2). In the past, ANN-based studies have shown R^2 values of 0.66 [8], 0.82 [62], 0.83 [63], 0.90 [53], and 0.94 [64]. Furthermore, the ANN algorithm has been used to perform different training functions (lbfgs, sgd, and ADAM) to predict the TBM performance, as shown in Figure 16. Comparing these results with the proposed study shows promising results for predicting TBM performance accurately. The third proposed model of this study is a model that integrates SVR and ANN. The integrated model also predicted accurate results from the dataset used for the previous two models. The correlation between predicted and actual TBM PR values for the SVR-ANN is plotted in Figure 15(c1,c2). These figures show that the ANN-based models provide higher accuracy because of their higher R^2 values. Recently, some optimized and integrated models were presented to predict the PR of TBM in different geological conditions and input parameters. The R^2 values of some of the studies are 0.905 for PSO-ANN, 0.912 for ICA-ANN [8], and 0.961 for PSO-ANN, 0.939 for ICA-ANN [7]. Comparing these results with the proposed model shows the practical applicability of the proposed model under different ground conditions.

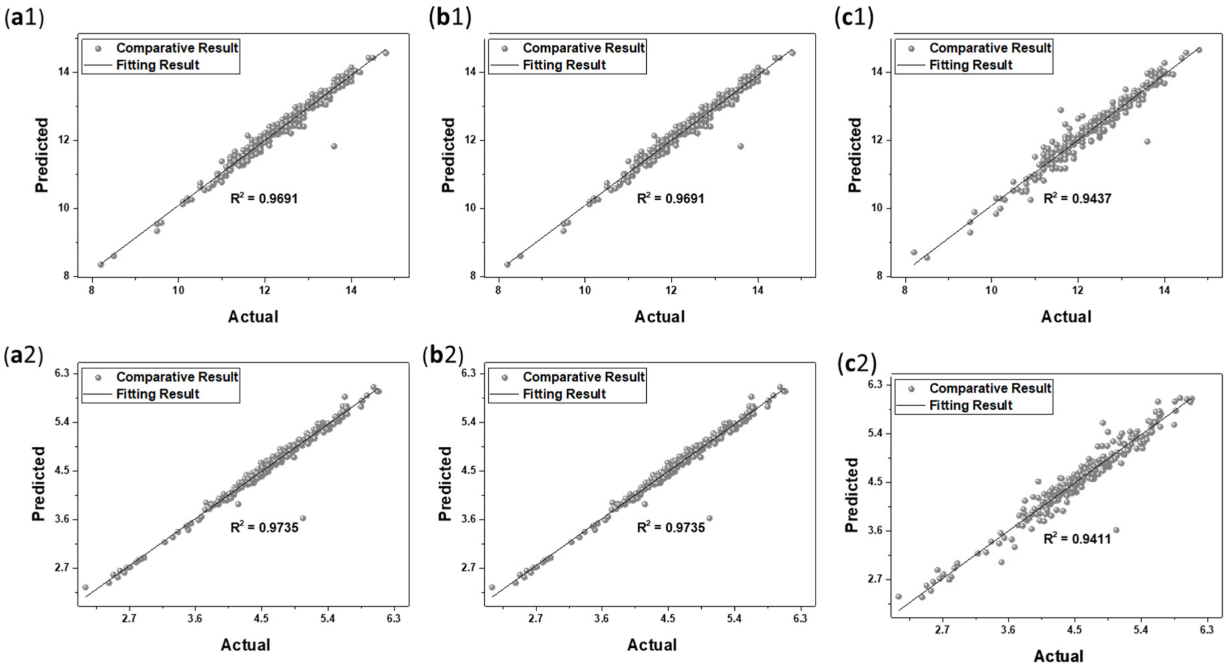

Figure 16. Regression analysis based on the actual and predicted TBM values; (**a1**,**b1**,**c1**) represents the penetration rate while (**a2**,**b2**,**c2**) represents the advanced rate for training functions lbfgs, sgd, and ADAM, respectively.

The performance indices RMSE, R^2, and VAF are evaluated and enlisted in Table 3. As shown in Table 3, compared with the ANN model, the SVM-ANN model provides better understanding. Additionally, a performance comparison of the SVM, ANN, and SVM-ANN models is shown in Figure 17. The VAF values are decimals for a comparative plot with R^2 and RSME. R^2 (0.9607, 0.9690, and 0.9694), RMSE (0.0436, 0.0343, and 0.0339), and VAF (96.07%, 96.90%, and 96.94%), are the performance predictors of the SVM, ANN, and SVM-ANN models, respectively. Significant improvement can be seen in the performance of these algorithms, respectively.

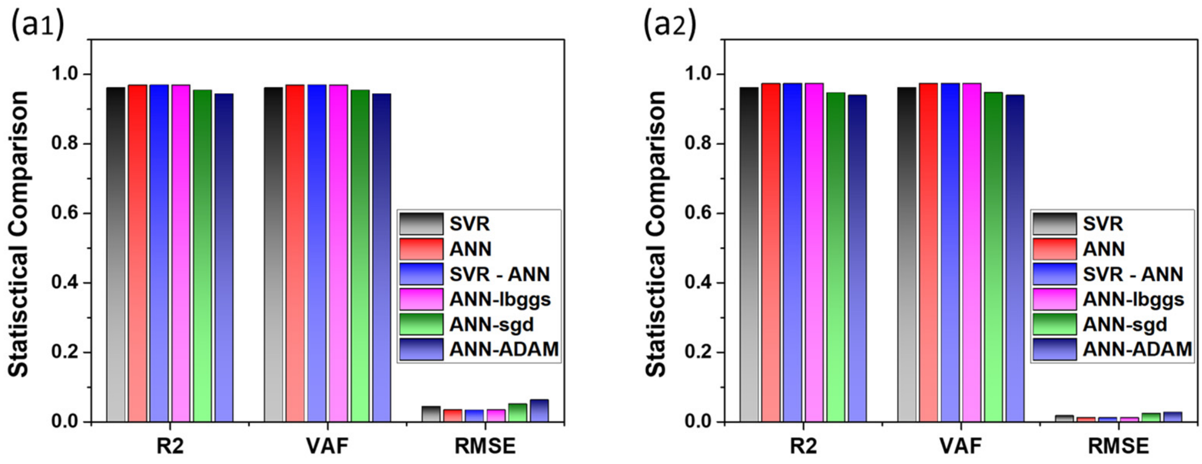

Figure 17. Regression analysis based on the actual and predicted TBM values: (**a1**) comparative results of models for the prediction of penetration rate; (**a2**) comparative results of models for the prediction of the advance rate.

A graphical user interface (GUI) for visualizing and analyzing the performance of a TBM, as shown in Figure 18, displays real-time data from the TBM. The data from the TBM are continuously monitored and analyzed through the AI, and predictions are carried out simultaneously. This shows the trends of the data generated from the AI through graphic representation in real time. The TBM GUI also provides the opportunity to analyze the data, display charts, and scatter plots for selected algorithms. It provides a user-friendly navigation environment to filter, sort, and select the desired algorithm and alert the system in the case of an undesirable event.

Table 3. Performance indices values for proposed algorithms.

Algorithm	PR			AR		
	R^2	RMSE	VAF	R^2	RMSE	VAF
SVR	0.960705	0.043628	0.960705	0.961474	0.017447	0.961782
ANN	0.969098	0.034308	0.969098	0.973454	0.012022	0.973659
SVR-ANN	0.9694	0.033973	0.969402	0.973602	0.011955	0.973843
ANN-lbfgs	0.969098	0.034308	0.969098	0.973454	0.012022	0.973659
ANN-sgd	0.953988	0.051084	0.954006	0.94675	0.024115	0.946851
ANN-ADAM	0.943473	0.062759	0.943655	0.939158	0.027553	0.939315

TBM DTs provide a virtual replica of a physical TBM. Using real-time data in a machine learning algorithm provides better performance monitoring in a virtual and remote environment. This helps the operator to identify issues, respond faster, and reduce TBM downtime. Thorough monitoring through the machine learning algorithm also helps to identify patterns that may lead to potential failure in advance. This helps with proactive maintenance. Furthermore, real-time data monitoring and TBM data analysis also provides the opportunity to identify optimal operating parameters. Reducing maintenance costs, minimizing downtime, improving safety, and increasing overall efficiency are the listed advantages of TBM digital twins.

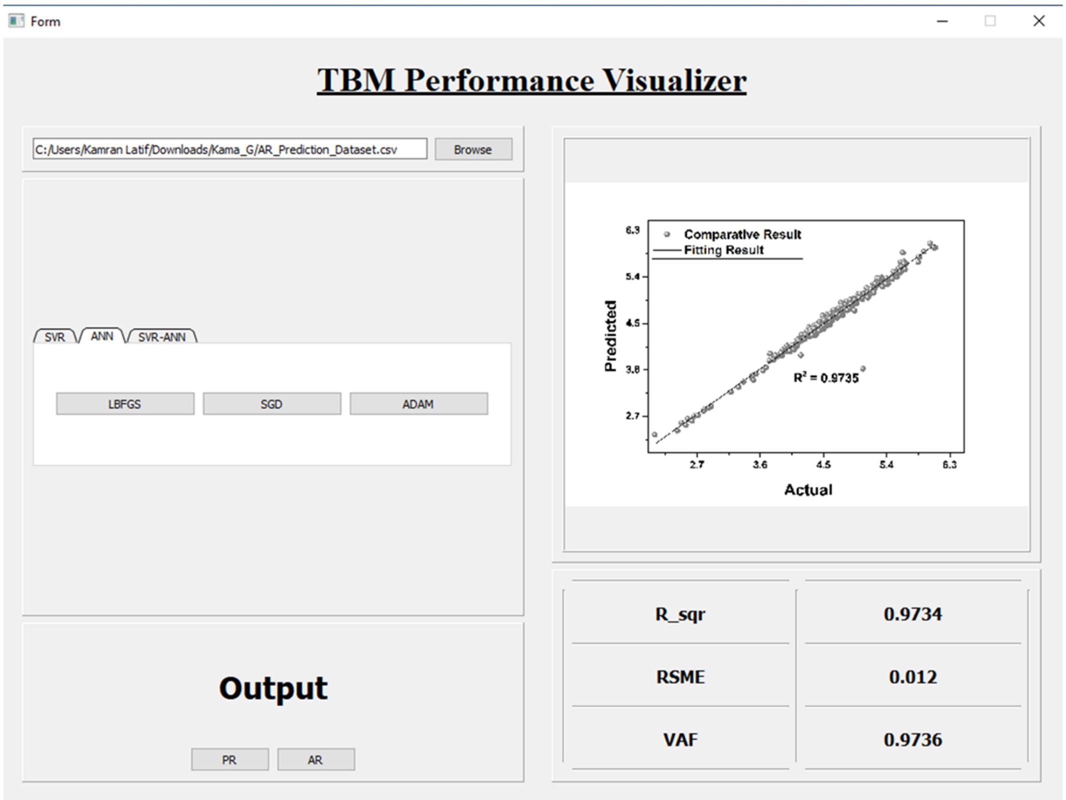

Figure 18. TBM performance visualizer.

6. Conclusions

Prediction of TBM performance is the most important factor for reliable cost estimation, project planning, and feasibility of mechanized excavation methods in tunneling projects. TBM tunneling is highly recommended for higher performance and better safety conditions than other tunneling methods. However, the heterogeneous and diverse data make it difficult to visualize the tunneling operation in its raw form. Visualizing TBM performance is critical for safe, effective, rapid, and risk-free tunnel construction according to the planned schedule. Therefore, the digital twin-driven framework for TBM performance prediction through machine learning, visualization, and monitoring is proposed for a better and easier way to visualize TBM performance. The proposed framework predicted the performance of the TBM operation through machine learning models and simultaneously visualized the results through digital twin modeling. For the theoretical demonstration of the framework, the data were collected from the open gripper hard rock TBM used at the Neelum Jhelum Hydroelectric Power (NJHEP) project tunnel. TBM digital twin modeling was performed based on the design drawing, design dimensions of electric and mechanical devices, CAD, and sensor-based data of the physical TBM. One thousand two hundred eighty-five (1285) data points of nine geotechnical and machine parameters, including boring energy, RPM, torque, thrust force, speed, gripper pressure, total revolutions, Q-value (theo), and Q-value$_{TBM}$, were considered for the development of these predictive models. SVR and ANN (BFGS quasi-Newton backpropagation algorithm)-based machine learning algorithms were used for TBM performance prediction. R^2 (0.9607, 0.9690, and 0.9694), RMSE (0.0436, 0.0343, and 0.0339), and VAF (96.07%, 96.90%, and 96.94%), were the SVM, ANN, and SVM-ANN model performance predictors, respectively, used for PR prediction. We conclude that the performance predictors (R^2, RSME, and VAF) showed higher accuracy in predicting TBM performance.

Furthermore, the predicted results of machine learning and operating parameters of the TBM were then linked with the TBM digital twin. Through AI scripting, the TBM RPM values were provided for the virtual TBM model, and TBM cutterhead rotation at different RPMs were recorded and visualized in a gaming platform. Digital twin modeling provides a visual demonstration of TBM performance to ease understanding for stakeholders. Based on the results, we conclude that the proposed digital twin-driven TBM framework has excellent potential to enhance the visual capabilities of construction performance, construction management, and decision-making. In terms of the limitations of the proposed framework, the digital twin framework lacks actual demonstration in a real environment with limited TBM access, considering the cost of the TBM. Secondly, for the theoretical demonstration of the framework, only rotational data were linked with the virtual TBM model. Future work may involve the thorough implementation of TBM DTs with ground interaction. Furthermore, the proposed framework considers data structure, and that interoperation of TBMs have a huge impact on the tunneling process, which will be covered in future studies.

Author Contributions: Conceptualization K.L. and A.S.; Methodology, K.L. and A.S.; Software, K.L.; Validation, K.L.; Formal Analysis, A.S.; Investigation, K.L. and A.S.; Resources, A.S. and J.S.; Data Curation, K.L. and A.S.; Writing—Original Draft Preparation, K.L.; Writing—Review and Editing, A.S. and J.S.; Visualization, K.L. and A.S.; Supervision, A.S. and J.S.; Project Administration, J.S.; Funding Acquisition, J.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Land, Infrastructure and Transport (National Research for Smart Construction Technology) (MSIT), Grant 20SMIP-A158708-01, National Research Foundation of Korea (NRF) grant funded by the Korean government (No. NRF-2019R1A2C2006577) and National Research Foundation of Korea (NRF) grant funded by the Korean government (MEST) (No. NRF 2018R1A5A1025137).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: Kamran Latif is extremely thankful to the Higher Education Commission of Pakistan for a scholarship for the Human Resource Development Initiative–Universities of Engineering Science and Technology.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

AI	Artificial intelligence
AEC	Architecture, engineering, and construction
ANFIS	Adaptive neuro–fuzzy interference system
ANN	Artificial neural network
AR	Advance rate
BFGS	Broyden–Fletcher–Goldfarb–Shanno algorithm
CAD	Computer-aided design
DT	Digital twin
FS	Feature space
GMDH	Group modeling of data handling
ICA	Imperialist competitive algorithm
IoT	Internet of things
ML	Machine learning
MLP	Multilayer perceptron
NJHEP	Neelum Jhelum hydroelectric project
NASA	National aeronautics and space administration
O&M	Operation and maintenance
PR	Penetration rate
PSO	Particle swarm optimization
ReLU	Rectified linear unit
RMR	Rock mass rating

R2	Coefficient of determination
RMSE	Root mean squared error
RPM	Revolutions per minute
RQD	Rock quality designation
SVM	Support vector machine
SVR	Support vector regression
TBM	Tunnel boring machine
UCS	Uniaxial compressive strength
UI	Utilization Index
VAF	Variance accounted for

References

1. Sharafat, A.; Khan, M.S.; Latif, K.; Seo, J. BIM-Based Tunnel Information Modeling Framework for Visualization, Management, and Simulation of Drill-and-Blast Tunneling Projects. *J. Comput. Civ. Eng.* **2021**, *35*, 04020068. [CrossRef]
2. Mahdevari, S.; Shahriar, K.; Yagiz, S.; Akbarpour Shirazi, M. A Support Vector Regression Model for Predicting Tunnel Boring Machine Penetration Rates. *Int. J. Rock Mech. Min. Sci.* **2014**, *72*, 214–229. [CrossRef]
3. Jung, J.-H.; Chung, H.; Kwon, Y.-S.; Lee, I.-M. An ANN to Predict Ground Condition Ahead of Tunnel Face Using TBM Operational Data. *KSCE J. Civ. Eng.* **2019**, *23*, 3200–3206. [CrossRef]
4. Zhou, J.; Qiu, Y.; Zhu, S.; Armaghani, D.J.; Li, C.; Nguyen, H.; Yagiz, S. Optimization of Support Vector Machine through the Use of Metaheuristic Algorithms in Forecasting TBM Advance Rate. *Eng. Appl. Artif. Intell.* **2021**, *97*, 104015. [CrossRef]
5. Sharafat, A.; Latif, K.; Park, S.; Seo, J. Risk Assessment of TBM Advanced Rate in the Difficult Ground Using Event Tree Analysis (ETA). In Proceedings of the KSCE, Virtual, 21–26 September 2021; pp. 39–40.
6. Sharafat, A.; Tanoli, W.A.; Raptis, G.; Seo, J.W. Controlled Blasting in Underground Construction: A Case Study of a Tunnel Plug Demolition in the Neelum Jhelum Hydroelectric Project. *Tunn. Undergr. Sp. Technol.* **2019**, *93*, 103098. [CrossRef]
7. Armaghani, D.J.; Koopialipoor, M.; Marto, A.; Yagiz, S. Application of Several Optimization Techniques for Estimating TBM Advance Rate in Granitic Rocks. *J. Rock Mech. Geotech. Eng.* **2019**, *11*, 779–789. [CrossRef]
8. Armaghani, D.J.; Mohamad, E.T.; Narayanasamy, M.S.; Narita, N.; Yagiz, S. Development of Hybrid Intelligent Models for Predicting TBM Penetration Rate in Hard Rock Condition. *Tunn. Undergr. Sp. Technol.* **2017**, *63*, 29–43. [CrossRef]
9. Latif, K.; Sharafat, A.; Seo, J. Predicting TBM Advanced Rate in Difficult Ground Conditions Based on the Artificial Neural Network Using Different Training Functions. In Proceedings of the KSCE, Virtual, 21–26 September 2021; pp. 23–24.
10. Latif, K.; Sharafat, A.; Park, S.; Seo, J. Digital Twin-Based Hybrid Approach to Visualize the Performance of TBM. In Proceedings of the KSCE, Busan, Republic of Korea, 20 October 2022; pp. 3–4.
11. Alvarez Grima, M.; Verhoef, P.N.W. Forecasting Rock Trencher Performance Using Fuzzy Logic. *Int. J. Rock Mech. Min. Sci.* **1999**, *36*, 413–432. [CrossRef]
12. Alvarez Grima, M.; Bruines, P.A.; Verhoef, P.N.W. Modeling Tunnel Boring Machine Performance by Neuro-Fuzzy Methods. *Tunn. Undergr. Sp. Technol.* **2000**, *15*, 259–269. [CrossRef]
13. Ghasemi, E.; Yagiz, S.; Ataei, M. Predicting Penetration Rate of Hard Rock Tunnel Boring Machine Using Fuzzy Logic. *Bull. Eng. Geol. Environ.* **2014**, *73*, 23–35. [CrossRef]
14. Yagiz, S.; Karahan, H. Application of Various Optimization Techniques and Comparison of Their Performances for Predicting TBM Penetration Rate in Rock Mass. *Int. J. Rock Mech. Min. Sci.* **2015**, *80*, 308–315. [CrossRef]
15. Jahed Armaghani, D.; Faradonbeh, R.S.; Momeni, E.; Fahimifar, A.; Tahir, M.M. Performance Prediction of Tunnel Boring Machine through Developing a Gene Expression Programming Equation. *Eng. Comput.* **2018**, *34*, 129–141. [CrossRef]
16. Benardos, A.G.; Kaliampakos, D.C. Modelling TBM Performance with Artificial Neural Networks. *Tunn. Undergr. Sp. Technol.* **2004**, *19*, 597–605. [CrossRef]
17. Zhou, C.; Ding, L.; Zhou, Y.; Zhang, H.; Skibniewski, M.J. Hybrid Support Vector Machine Optimization Model for Prediction of Energy Consumption of Cutter Head Drives in Shield Tunneling. *J. Comput. Civ. Eng.* **2019**, *33*, 04019019. [CrossRef]
18. Afradi, A.; Ebrahimabadi, A.; Hallajian, T. Prediction of the Penetration Rate and Number of Consumed Disc Cutters of Tunnel Boring Machines (TBMs) Using Artificial Neural Network (ANN) and Support Vector Machine (SVM)—Case Study: Beheshtabad Water Conveyance Tunnel in Iran. *Asian J. Water Environ. Pollut.* **2019**, *16*, 49–57. [CrossRef]
19. Afradi, A.; Ebrahimabadi, A.; Hallajian, T. Prediction of Tunnel Boring Machine Penetration Rate Using Ant Colony Optimization, Bee Colony Optimization and the Particle Swarm Optimization, Case Study: Sabzkooh Water Conveyance Tunnel. *Min. Miner. Depos.* **2020**, *14*, 75–84. [CrossRef]
20. Yu, G.; Wang, Y.; Mao, Z.; Hu, M.; Sugumaran, V.; Wang, Y.K. A Digital Twin-Based Decision Analysis Framework for Operation and Maintenance of Tunnels. *Tunn. Undergr. Sp. Technol.* **2021**, *116*, 104125. [CrossRef]
21. Bonitz, L.; Pugachev, A.; Arnold, M.; Janoske, U.; Hassfeld, S.; Abel, D.; Bicsak, A.; Mueller, C. "Digital Twin" Based Approach to Patient Specific Diagnosis and Therapy of OSA. *Sleep Med.* **2019**, *64*, S45. [CrossRef]
22. Björnsson, B.; Borrebaeck, C.; Elander, N.; Gasslander, T.; Gawel, D.R.; Gustafsson, M.; Jörnsten, R.; Lee, E.J.; Li, X.; Lilja, S.; et al. Digital Twins to Personalize Medicine. *Genome Med.* **2020**, *12*, 4. [CrossRef]

23. Terkaj, W.; Gaboardi, P.; Trevisan, C.; Tolio, T.; Urgo, M. A Digital Factory Platform for the Design of Roll Shop Plants. *CIRP J. Manuf. Sci. Technol.* **2019**, *26*, 88–93. [CrossRef]
24. Urbina Coronado, P.D.; Lynn, R.; Louhichi, W.; Parto, M.; Wescoat, E.; Kurfess, T. Part Data Integration in the Shop Floor Digital Twin: Mobile and Cloud Technologies to Enable a Manufacturing Execution System. *J. Manuf. Syst.* **2018**, *48*, 25–33. [CrossRef]
25. Zhuang, C.; Liu, J.; Xiong, H. Digital Twin-Based Smart Production Management and Control Framework for the Complex Product Assembly Shop-Floor. *Int. J. Adv. Manuf. Technol.* **2018**, *96*, 1149–1163. [CrossRef]
26. Bilberg, A.; Malik, A.A. Digital Twin Driven Human–Robot Collaborative Assembly. *CIRP Ann.* **2019**, *68*, 499–502. [CrossRef]
27. Coraddu, A.; Oneto, L.; Baldi, F.; Cipollini, F.; Atlar, M.; Savio, S. Data-Driven Ship Digital Twin for Estimating the Speed Loss Caused by the Marine Fouling. *Ocean Eng.* **2019**, *186*, 106063. [CrossRef]
28. Cai, Y.; Wang, Y.; Burnett, M. Using Augmented Reality to Build Digital Twin for Reconfigurable Additive Manufacturing System. *J. Manuf. Syst.* **2020**, *56*, 598–604. [CrossRef]
29. Xiong, M.; Wang, H.; Fu, Q.; Xu, Y. Digital Twin–Driven Aero-Engine Intelligent Predictive Maintenance. *Int. J. Adv. Manuf. Technol.* **2021**, *114*, 3751–3761. [CrossRef]
30. Xu, Z.; Ji, F.; Ding, S.; Zhao, Y.; Zhou, Y.; Zhang, Q.; Du, F. Digital Twin-Driven Optimization of Gas Exchange System of 2-Stroke Heavy Fuel Aircraft Engine. *J. Manuf. Syst.* **2021**, *58*, 132–145. [CrossRef]
31. Zheng, Y.; Yang, S.; Cheng, H. An Application Framework of Digital Twin and Its Case Study. *J. Ambient. Intell. Humaniz. Comput.* **2019**, *10*, 1141–1153. [CrossRef]
32. Sharafat, A.; Latif, K.; Park, S.; Seo, J. Digital Twin-Driven Optimization of Blast Design for Underground Construction. In Proceedings of the KSCE, Seoul, Republic of Korea, 20 October 2022; pp. 1–2.
33. Kim, J.; Kim, S.-A. Lifespan Prediction Technique for Digital Twin-Based Noise Barrier Tunnels. *Sustainability* **2020**, *12*, 2940. [CrossRef]
34. Yu, G.; Zhang, S.; Hu, M.; Wang, Y.K. Prediction of Highway Tunnel Pavement Performance Based on Digital Twin and Multiple Time Series Stacking. *Adv. Civ. Eng.* **2020**, *2020*, 8824135. [CrossRef]
35. Sharafat, A.; Khan, M.S.; Latif, K.; Tanoli, W.A.; Park, W.; Seo, J. BIM-GIS-Based Integrated Framework for Underground Utility Management System for Earthwork Operations. *Appl. Sci.* **2021**, *11*, 5721. [CrossRef]
36. Qi, Q.; Tao, F.; Hu, T.; Anwer, N.; Liu, A.; Wei, Y.; Wang, L.; Nee, A.Y.C. Enabling Technologies and Tools for Digital Twin. *J. Manuf. Syst.* **2021**, *58*, 3–21. [CrossRef]
37. Boje, C.; Guerriero, A.; Kubicki, S.; Rezgui, Y. Towards a Semantic Construction Digital Twin: Directions for Future Research. *Autom. Constr.* **2020**, *114*, 103179. [CrossRef]
38. Pan, Y.; Zhang, L. A BIM-Data Mining Integrated Digital Twin Framework for Advanced Project Management. *Autom. Constr.* **2021**, *124*, 103564. [CrossRef]
39. Grieves, M. Digital Twin: Manufacturing Excellence through Virtual Factory Replication. *White Pap.* **2014**, *1*, 1–7.
40. Richnák, P. Current Trend of Industry 4.0 in Logistics and Transformation of Logistics Processes Using Digital Technologies: An Empirical Study in the Slovak Republic. *Logistics* **2022**, *6*, 79. [CrossRef]
41. Pech, M.; Vrchota, J. Classification of Small- and Medium-Sized Enterprises Based on the Level of Industry 4.0 Implementation. *Appl. Sci.* **2020**, *10*, 5150. [CrossRef]
42. Tubis, A.A.; Grzybowska, K. In Search of Industry 4.0 and Logistics 4.0 in Small-Medium Enterprises—A State of the Art Review. *Energies* **2022**, *15*, 8595. [CrossRef]
43. Glaessgen, E.; Stargel, D. The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles. In Proceedings of the 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, Honolulu, HI, USA, 23–26 April 2012.
44. Tao, F.; Zhang, M. Digital Twin Shop-Floor: A New Shop-Floor Paradigm Towards Smart Manufacturing. *IEEE Access* **2017**, *5*, 20418–20427. [CrossRef]
45. Alexopoulos, K.; Nikolakis, N.; Chryssolouris, G. Digital Twin-Driven Supervised Machine Learning for the Development of Artificial Intelligence Applications in Manufacturing. *Int. J. Comput. Integr. Manuf.* **2020**, *33*, 429–439. [CrossRef]
46. Koopialipoor, M.; Nikouei, S.S.; Marto, A.; Fahimifar, A.; Jahed Armaghani, D.; Mohamad, E.T. Predicting Tunnel Boring Machine Performance through a New Model Based on the Group Method of Data Handling. *Bull. Eng. Geol. Environ.* **2019**, *78*, 3799–3813. [CrossRef]
47. Gordan, B.; Koopialipoor, M.; Clementking, A.; Tootoonchi, H.; Tonnizam Mohamad, E. Estimating and Optimizing Safety Factors of Retaining Wall through Neural Network and Bee Colony Techniques. *Eng. Comput.* **2019**, *35*, 945–954. [CrossRef]
48. Min, Q.; Lu, Y.; Liu, Z.; Su, C.; Wang, B. Machine Learning Based Digital Twin Framework for Production Optimization in Petrochemical Industry. *Int. J. Inf. Manag.* **2019**, *49*, 502–519. [CrossRef]
49. Tao, F.; Sui, F.; Liu, A.; Qi, Q.; Zhang, M.; Song, B.; Guo, Z.; Lu, S.C.-Y.; Nee, A.Y.C. Digital Twin-Driven Product Design Framework. *Int. J. Prod. Res.* **2019**, *57*, 3935–3953. [CrossRef]
50. Umair, M.; Sharafat, A.; Lee, D.-E.; Seo, J. Impact of Virtual Reality-Based Design Review System on User's Performance and Cognitive Behavior for Building Design Review Tasks. *Appl. Sci.* **2022**, *12*, 7249. [CrossRef]
51. Gao, Y.; Qian, S.; Li, Z.; Wang, P.; Wang, F.; He, Q. Digital Twin and Its Application in Transportation Infrastructure. In Proceedings of the IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI), Beijing, China, 15 July–15 August 2021; pp. 298–301. [CrossRef]

52. Huang, X.; Liu, Q.; Liu, H.; Zhang, P.; Pan, S.; Zhang, X.; Fang, J. Development and In-Situ Application of a Real-Time Monitoring System for the Interaction between TBM and Surrounding Rock. *Tunn. Undergr. Sp. Technol.* **2018**, *81*, 187–208. [CrossRef]
53. Yagiz, S.; Gokceoglu, C.; Sezer, E.; Iplikci, S. Application of Two Non-Linear Prediction Tools to the Estimation of Tunnel Boring Machine Performance. *Eng. Appl. Artif. Intell.* **2009**, *22*, 808–814. [CrossRef]
54. Barton, N. TBM Performance Estimation in Rock Using Q(TBM). *Tunnels Tunn. Int.* **1999**, *31*, 30–34.
55. Gong, Q.M.; Zhao, J. Development of a Rock Mass Characteristics Model for TBM Penetration Rate Prediction. *Int. J. Rock Mech. Min. Sci.* **2009**, *46*, 8–18. [CrossRef]
56. Barton, N.; Lien, R.; Lunde, J. Engineering classification of rock masses for the design of tunnel support. *Rock Mech.* **1974**, *6*, 189–236. [CrossRef]
57. Naithani, A.K. Rock Mass Classification and Support Design using the Q-System. *J. Geol. Soc. India* **2019**, *94*, 443. [CrossRef]
58. Shi, X.; Zhou, J.; Wu, B.; Huang, D.; Wei, W. Support Vector Machines Approach to Mean Particle Size of Rock Fragmentation Due to Bench Blasting Prediction. *Trans. Nonferrous Met. Soc. China* **2012**, *22*, 432–441. [CrossRef]
59. Adoko, A.-C.; Jiao, Y.-Y.; Wu, L.; Wang, H.; Wang, Z.-H. Predicting Tunnel Convergence Using Multivariate Adaptive Regression Spline and Artificial Neural Network. *Tunn. Undergr. Sp. Technol.* **2013**, *38*, 368–376. [CrossRef]
60. Mahdevari, S.; Torabi, S.R. Prediction of Tunnel Convergence Using Artificial Neural Networks. *Tunn. Undergr. Sp. Technol.* **2012**, *28*, 218–228. [CrossRef]
61. Sharafat, A.; Latif, K.; Seo, J. Risk Analysis of TBM Tunneling Projects Based on Generic Bow-Tie Risk Analysis Approach in Difficult Ground Conditions. *Tunn. Undergr. Sp. Technol.* **2021**, *111*, 103860. [CrossRef]
62. Yavari, M.; Mahdavi, S. Prediction of Penetration Rate of TBM Using ANN. In Proceedings of the National Mining Conference, Breckenridge, CO, USA, 19–23 June 2005; pp. 1–10.
63. Salimi, A.; Esmaeili, M. Utilising of Linear and Non-Linear Prediction Tools for Evaluation of Penetration Rate of Tunnel Boring Machine in Hard Rock Condition. *Int. J. Min. Miner. Eng.* **2013**, *4*, 249. [CrossRef]
64. Javad, G.; Narges, T. Application of Artificial Neural Networks to the Prediction of Tunnel Boring Machine Penetration Rate. *Min. Sci. Technol.* **2010**, *20*, 727–733. [CrossRef]