DATA-DRIVEN TECHNIQUES FOR REAL-TIME SAFETY MANAGEMENT IN TUNNEL ENGINEERING USING TBM DATA

Mohanarangan Veerappermal Devarajan,
Cambria Solutions Sacramento
gc4mohan@gmail.com

Aceng Sambas,
Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Campus Besut, 22200 Terengganu, Malaysia
Department of Mechanical Engineering, Universitas Muhammadiyah Tasikmalaya, Tamansari Gobras 46196 Tasikmalaya, Indonesia
aceng.sambs@gmail.com

ABSTRACT

Tunnel engineering entails large hazards, protracted building timeframes, and expensive prices. Because they gather so much monitoring data, tunnel boring machines (TBMs) are essential to increasing efficiency and safety. In order to improve safety management and decision-making, this study presents a hybrid data mining technique for automating real-time TBM data processing. The system analyses TBM parameters, finds anomalies, categorises geological formations, and forecasts the rate of penetration (ROP) by integrating association rule mining, decision tree classification, and neural network models. When used on a tunnel project in China, this method greatly improved operational efficiency and safety management by analysing TBM data with high precision and efficacy.

Keywords: Tunnel Boring Machines, real-time data analysis, hybrid data mining, association rule mining, decision tree classification, neural network modeling.

1 INTRODUCTION

Tunnel engineering is an enormous project with high construction costs, lengthy construction schedules, and considerable dangers. In this operation, tunnel boring machines (TBMs) are essential for increasing safety and efficiency. TBMs gather copious amounts of monitoring data during tunnelling in order to guarantee safety. Effective real-time data analysis is still difficult to do, frequently involving human labour that could be dangerous. In order to automatically process real-time TBM data, this research presents a hybrid data mining (DM) technique. Our goal is to improve mining efficiency and help safety management by integrating three distinct DM approaches. We investigate correlations among TBM parameters using an association rule approach to facilitate on-site anomaly detection. To help with building decision-making, we also use a decision tree model to categorise formation data. To identify irregularities and issue early warnings, neural network models lastly assess the rate of penetration (ROP). Its accuracy and effectiveness in analysing TBM monitoring data for safety management were shown when this technology was applied to a Chinese tunnel project.

The ability of TBMs to manage large-scale, expensive, and time-consuming projects—many of which involve intricate technical risks and challenges—makes them indispensable for tunnel building. A wide range of operational characteristics, including stress, current, flow, and gas pressure, are recorded by the multiple sensors and monitoring instruments that are installed in modern TBMs. These data lower construction hazards, improve efficiency, and guarantee security. The creation of cyber-physical systems that enable automated data gathering, storage, and visualisation is a result of advancements in sensing technologies. For tunnel building, manual data analysis is still dangerous and inconsistent even with these developments. In February 2018, for example, a severe water inrush accident occurred during the construction of a Guangdong tunnel, resulting in over 10 fatalities and significant financial damages. The TBM came across a porous layer that did not exist during the geological study, which is what caused the accident. Efficient and precise data analysis could help avert such occurrences. However, there are a number of issues with TBM monitoring data that make real-time automatic data processing difficult.

Difficulty in identifying essential variables: To reduce risk, data from multiple subsystems must be recorded during tunnel building. It is difficult to manually grasp and predict the interactions between more than a hundred monitoring metrics. It is challenging to find all pertinent variables for precise data analysis because of this intricacy. Error as a result of uneven soil: Unlike above-ground constructions, the construction of underground tunnels requires precise geological information. The most common method for collecting geological data is borehole sampling, from which formations between sampling locations are calculated using linear fitting. These estimates frequently deviate from the real circumstances, which lowers data accuracy and may cause weak formations that could present unforeseen problems during construction to be overlooked. Complicated nonlinear relationships between monitoring variables: During TBM operation, the interactions between the machinery, shield, and soil produce complicated correlations between quantities that are challenging to represent with straightforward mathematical models. When it comes to precisely analysing these interactions, traditional statistical methods frequently fall short.

- Provide a data mining-based system that automatically evaluates real-time TBM data to improve tunnel construction safety.
- Use algorithms to analyse TBM data and identify patterns that could lead to early problem detection.
- Classify and comprehend the geological formations that are discovered while tunnelling with the help of decision trees.
- To track the rate of penetration (ROP) and identify any irregularities promptly, use neural networks.
- Verify the accuracy and effectiveness of this novel strategy by testing it on a real tunnel project.

The existing approaches to real-time data analysis from TBMs are insufficient since they mostly depend on labor-intensive, erroneous, and slow manual operations. The vast volumes and complexity of data created during tunnel construction remain unmanageable for automated systems, despite technological advancements in data collection. This disparity emphasises the necessity for a more advanced, automated strategy that incorporates several data mining methods in order to increase tunnel engineering efficiency and safety.

Tunnel building is a complicated, high-risk procedure that uses TBMs to produce enormous volumes of data. A large portion of this data is now analysed manually, which can result in errors and even pose a risk to public safety. The goal of this project is to develop a hybrid data mining system that automatically processes and interprets TBM data by combining several analytical techniques. Enhancing safety management and decision-making throughout tunnel construction is intended to lower the risk of accidents and boost overall productivity.

2 LITERATURE SURVEY

Tunnel Boring Machines (TBMs) are used in tunnel construction, and Leng (2020) describe a hybrid data mining system that improves real-time safety management. For automatic TBM data analysis, this system integrates neural network models, decision tree categorization, and association rule mining. Applying it to a Chinese tunnel project, it demonstrated high accuracy in predicting rate of penetration (ROP) and classifying geological formations (86%), as well as anomaly identification (95%), greatly increasing productivity and safety. The method has the potential to enhance safety management, as seen by its effectiveness in this project. In the future, the method may be applied to different tunnelling projects to further increase accuracy and efficiency. Other potential upgrades would involve incorporating IoT technology and real-time data systems.

Using data from Tunnel Boring Machines (TBMs), Zhao (2019) present a data-driven framework to anticipate geological types in tunnel building. This system enhances real-time safety monitoring and decision-making by automating TBM data processing through the integration of neural network models, decision tree categorization, and association rule mining. When the framework was used for a tunnel project in China, it demonstrated great accuracy in identifying anomalies (95%) and categorising geological formations (86%). The BP neural network's R2 value was 0.92 when it came to forecasting the rate of penetration (ROP). The safety and operational efficiency of this hybrid system were greatly increased. To further increase accuracy and efficiency, future improvements can include incorporating IoT technologies, real-time data platforms, and extending the system's use to different tunnelling projects.

The safety and efficiency of Tunnel Boring Machine (TBM) construction can be enhanced by the use of information and intelligence technologies, as examined by Li (2019). They examine real-time data from TBM sensors using sophisticated data mining techniques like decision tree classification, association rule mining, and neural network models. Accurate, in-the-moment decision-making is supported by this ongoing data collection. Critical metrics including anomaly detection, geological formation categorization, and penetration rate (ROP) can be predicted with the use of predictive model building. Because proactive management and accident prevention are made possible by early detection of potential risks and anomalies, this technique improves safety. Furthermore, by automating data processing and so decreasing the need for human processes, it increases operational efficiency. To further increase safety and efficiency, the study recommends expanding applications across different tunnelling projects and integrating real-time data platforms and IoT technologies in the future.

Xue (2019) use big data to forecast unfavourable geological conditions and issue early warnings for TBM tunnelling. The goal of the project is to improve operational effectiveness and safety by utilising real-time data from TBM sensors and sophisticated analytics. Using massive amounts of data from TBM operations, big data analytics may detect and forecast unfavourable geological conditions. Prompt decision-making is made possible by ongoing observation and real-time data processing. Because they foresee and reduce dangers, predictive models improve safety by forecasting geological issues and issuing early warnings. This proactive strategy lowers downtime and takes early action to alleviate geological problems, which boosts efficiency. Before possible risks have an impact on tunnelling operations, early warning systems notify operators of them.

Marcher (2020) examine the advantages and disadvantages of using machine learning in tunnelling. Through hazard prediction and identification, operational optimisation to minimise downtime, and predictive maintenance to avert equipment failures, machine learning can improve tunnel construction safety. The study does, however, also highlight certain difficulties, like obtaining reliable data from TBM sensors, creating precise models that can manage intricate tunnelling conditions, and incorporating these fixes with already-in-place systems. Notwithstanding these challenges, continued developments in data analytics and machine learning provide more innovations in tunnel management and construction.

Data-driven strategies were employed in a study by Wang (2019) to predict and rectify axis attitude variations during shield tunnelling. Their objectives were to enhance the forecasting of deviations and create an automated remedy system. The research employed real-time monitoring systems and machine learning techniques to collect and evaluate operational data, such as soil conditions and machine position. Using both historical and current data, a predictive model was developed and verified. These forecasts were subsequently used to inform correction procedures that were incorporated into the tunnelling control system for automated modifications. The data-driven strategy led to more accurate tunnelling operations by demonstrating a considerable improvement in deviation forecast accuracy and rectification efficiency. According to the study's findings, this strategy provides a reliable means of regulating deviations; going forward, efforts will be focused on improving the models and integrating cutting-edge technology.

Shen (2014) suggest integrating machine guidance and operations monitoring as part of an integrated strategy to improve tunnel construction. Their goal is to increase tunneling's accuracy and efficiency. Tunnelling equipment are precisely guided by automated controls and GPS, among other advanced navigation systems. Real-time monitoring systems use sensors to collect precise data while tracking operational parameters including soil conditions and machine position. A thorough understanding of tunnelling operations is obtained through the integration of various systems, facilitating prompt and well-informed modifications and decision-making. The outcomes demonstrate notable gains in precision, effectiveness, and flexibility to shifting circumstances, which minimises downtime. This methodology offers a strong foundation for subsequent developments in tunnel building, emphasising the enhancement of sophisticated technology and the optimisation of current systems.

Tunnel boring machines (TBMs) were evaluated and improved by Zhang (2016) through the use of simulation-based analysis. They discovered important elements that influence TBM performance, such as machine settings and soil conditions, by building a virtual model using actual data. This method lowers expenses, increases efficiency, and optimises processes. Subsequent studies will enhance these models and broaden their use.

The application of artificial neural networks (ANNs) in tunnel engineering turned out the subject of a thorough review by Wang (2020). They looked at a number of applications, including optimising construction procedures and forecasting ground behaviour. In order to determine the advantages and disadvantages of ANN models over conventional techniques, the study examined ANN models, data sources, and performance indicators across tunnelling disciplines. They discovered ANNs improve decision-making in stability analysis and tunnel construction, implying that by enhancing predictive capacities and enabling real-time monitoring systems, ANNs could completely transform the tunnel engineering industry. The goals of future research are to improve ANN models, broaden their applications, and solve issues with interpretability and data quality.

Liu (2015) examined schedule risks in TBM tunnelling using Adaptive CYCLONE simulation, paying particular attention to managing geological uncertainties for improved scheduling precision. Their method evaluated risk scenarios to optimise TBM operations and reduce delays, and it incorporated real-time data to modify schedules in response to shifting geological conditions. The simulation demonstrated promise to improve project management by integrating geologically informed scheduling and enhancing decision-making in tunnelling projects, since the study indicated that it enhanced the forecast of project delays and interruptions. Subsequent investigations endeavour to enhance the models and expand their utilisation in other tunnelling scenarios.

In order to analyse Tunnel Boring Machine (TBM) activities, Song (2019) provide a novel fuzzy c-means clustering technique for segmenting time series data. By efficiently utilising TBM sensor data, their method improves precision in identifying phases like as cutting and advancing. When compared to conventional approaches, this method provides insights about TBM performance patterns and efficiency while improving accuracy. It can optimise operating strategies and maintenance schedules through its applications in real-time monitoring and decision support systems. The study finds that while this approach is useful for comprehending TBM operations, further research is needed to improve and implement it in other tunnelling contexts and machine kinds.

Tunnel Boring Machine (TBM) performance characteristics were predicted in real-time by Nagrecha (2020) using Recurrent Neural Networks (RNNs). They discovered that RNNs outperformed conventional techniques in accurately predicting measures such as advance rate and cutterhead torque. This method facilitates proactive decision-making in tunnelling projects by providing insights into the variables influencing TBM productivity and efficiency. To further improve efficiency and performance prediction, the study recommends improving RNN models for wider use across various tunnelling circumstances and types of TBM.

Basava Ramanjaneyulu Gudivaka (2019) investigated the application of big data approaches, particularly Hadoop, to forecast silicon content in blast furnace smelting operations. The study emphasises the integration of production records, sensor data, and environmental variables with big data analytics to enhance accuracy and efficiency compared to traditional empirical methods. The Hadoop framework facilitates real-time monitoring, predictive maintenance, and optimised furnace operations, hence improving steel quality and minimising downtime. Notwithstanding obstacles in data integration and real-time processing, the study underscores the promise of big data in process optimisation and predictive maintenance via collaborative initiatives.

Sareddy (2020) examined the impact of employee engagement tactics and compensation on retention across Pakistan's industrial and service sectors. The analysis of data from 1,054 employees revealed that both direct (consultative and delegative) and indirect (worker unions and directors) involvement tactics are critical retention determinants, with delegative participation exerting the greatest influence. Compensation was identified as a moderating factor in this connection, improving retention when engagement tactics are successful. The study emphasises the significance of customised engagement and compensation strategies for enhancing employee retention, especially in developing countries.

Allur (2020) examined the utilisation of Big Data, Decision Support Systems (DSS), and Mixed-Integer Linear Programming (MILP) in enhancing agricultural supply chain management (ASCM). The research emphasises the ways in which data-driven methodologies improve resource distribution, scheduling, and performance indicators, such as cost efficiency, dependability, and precision. The incorporation of these technologies in ASCM demonstrates considerable potential for enduring efficiency improvements and sustainable agriculture methods. The study highlights the significance of utilising sophisticated computational models to enhance decision-making and operational results in agriculture.

Kodadi (2020) presented a hybrid architecture for cloud computing security that combines the Immune Cloning Algorithm with Data-Driven Threat Mitigation (d-TM). The program, inspired by biological immune systems, improves danger identification, minimises false positives, and expedites reaction times. Simulations attained a 93% detection rate, 5% false positive rate, and a response time of 120 ms, surpassing traditional methods such as CSA and NLP. The research underscores the architecture's scalability, adaptability, and cost-effectiveness, while also noting possible enhancements to edge and quantum computing for augmented security.

Allur (2021) investigated resource allocation optimisation in cloud data centres with the introduction of a novel load-balancing methodology. The proposed solution utilises edge computing, artificial intelligence, and machine learning to rectify the deficiencies of conventional methodologies in dynamic cloud environments. This method improves scalability, efficiency, and system responsiveness by strategically allocating workloads among data centres and virtual machines. This research underscores the necessity for adaptive methodologies to enhance resource utilisation and performance, hence advancing cloud data centre management.

Allur (2021) examined sophisticated load-balancing techniques for enhancing resource distribution in cloud data centres. The research underscores the shortcomings of conventional techniques in dynamic cloud settings and suggests a novel strategy utilising edge computing, artificial intelligence, and machine learning. This method improves scalability, system efficiency, and responsiveness by judiciously allocating workloads among virtual machines and data centres. The study highlights the necessity for adaptive solutions to enhance resource utilisation and tackle issues in contemporary cloud architecture.

Multivariate Adaptive Regression Splines (MARS), Softmax Regression, and Histogram-Based Gradient Boosting are some of the most complex statistical and machine learning techniques applied by Narla et al. (2021) in analyzing the predictiveness of healthcare modeling within the cloud computing environment. The researchers found that integrating the three algorithms is significant to provide higher computational performance and better prediction accuracy, depending on the large size of healthcare data. The proposed model solves the problem of scaling and real-time analysis and hence has vast applications in predictive analytics of cloud-based medical system.

Peddi et al. (2018) demonstrate advancement in geriatric care by using machine learning algorithms and artificial intelligence in applications to predict the risks of senior patients developing dysphagia, delirium, and falls. Their work focuses on using predictive models to enhance early diagnosis and intervention tactics in improving patient safety and results. This paper addresses the growing need for data-driven strategies in the management of issues in geriatric health with the use of sophisticated computational methodologies.

Peddi et al. (2019) explored the use of AI and machine learning algorithms in fall prevention, management of chronic diseases, and the prediction of applications in healthcare, specifically in geriatric care. Their study puts forward how recent computational methods can be integrated into early intervention tactics, increase patient safety, and improve health outcomes. This paper demonstrates how predictive analytics can transform geriatric healthcare by underlining its importance in meeting the wide-ranging demands of elderly people.

Valivarthi et al. (2021) discuss the integration of cloud computing and artificial intelligence techniques to create advanced healthcare prediction models. To enhance the accuracy and efficiency of predictions, the study employs ABC-ANFIS (Artificial Bee Colony with Adaptive Neuro-Fuzzy Inference System) and BBO-FLC (Biogeography-Based Optimisation with Fuzzy Logic Control). This research discusses how evolutionary algorithms and fuzzy logic systems can be integrated to solve complex healthcare problems and improve cloud-based decision-making.

Narla et al. (2019) make use of the LSTM networks along with ACO for disease prediction and explore cloud computing in combination with healthcare, showing how this method combines predictive modelling power of LSTM with optimisation skills of ACO to get improved accuracy as well as scalability. This innovative approach addresses problems in healthcare, as it has enabled proactive health management and appropriate disease prediction under cloud-based setups.

Multivariate Adaptive Regression Splines (MARS), Softmax Regression, and Histogram-Based Gradient Boosting are some of the most complex statistical and machine learning techniques applied by Narla et al. (2021) in analyzing the predictiveness of healthcare modeling within the cloud computing environment. The researchers found that integrating the three algorithms is significant to provide higher computational performance and better prediction accuracy, depending on the large size of healthcare data. The proposed model solves the problem of scaling and real-time analysis and hence has vast applications in predictive analytics of cloud-based medical system.

Peddi et al. (2018) demonstrate advancement in geriatric care by using machine learning algorithms and artificial intelligence in applications to predict the risks of senior patients developing dysphagia, delirium, and falls. Their work focuses on using predictive models to enhance early diagnosis and intervention tactics in improving patient safety and results. This paper addresses the growing need for data-driven strategies in the management of issues in geriatric health with the use of sophisticated computational methodologies.

Peddi et al. (2019) explored the use of AI and machine learning algorithms in fall prevention, management of chronic diseases, and the prediction of applications in healthcare, specifically in geriatric care. Their study puts forward how recent computational methods can be integrated into early intervention tactics, increase patient safety, and improve health outcomes. This paper demonstrates how predictive analytics can transform geriatric healthcare by underlining its importance in meeting the wide-ranging demands of elderly people.

Valivarthi et al. (2021) discuss the integration of cloud computing and artificial intelligence techniques to create advanced healthcare prediction models. To enhance the accuracy and efficiency of predictions, the study employs ABC-ANFIS (Artificial Bee Colony with Adaptive Neuro-Fuzzy Inference System) and BBO-FLC (Biogeography-Based Optimisation with Fuzzy Logic Control). This research discusses how evolutionary algorithms and fuzzy logic systems can be integrated to solve complex healthcare problems and improve cloud-based decision-making.

Narla et al. (2019) make use of the LSTM networks along with ACO for disease prediction and explore cloud computing in combination with healthcare, showing how this method combines predictive modelling power of LSTM with optimisation skills of ACO to get improved accuracy as well as scalability. This innovative approach addresses problems in healthcare, as it has enabled proactive health management and appropriate disease prediction under cloud-based setups.

A cloud-integrated Smart Healthcare Framework by Narla et al. (2019) uses LightGBM for fast data processing, multinomial logistic regression for health risk analysis, and self-organising maps (SOMs) for data patterns. Scalable and real-time, the system stores and analyses data to improve healthcare decision-making. With a 95% AUC, the framework surpasses standard models in accuracy and recall, recognising health hazards and enabling personalised patient care. By using powerful machine learning, it allows immediate interventions and improves healthcare results through accurate and individualised treatment regimens.

3 METHODOLOGY

To improve safety management and decision-making during tunnel building, this project focuses on creating a hybrid data mining system to evaluate real-time data from tunnel boring machines (TBMs). Association rule mining, decision tree classification, and neural network modelling are the three primary components that the system integrates.

3.1 Association Rule Mining for Anomaly Detection

Using association rule mining, the initial stage is to identify correlations between TBM operational metrics that can point to possible abnormalities during tunnel building. Various operational data, including stress, current, flow, and gas pressure, are recorded by the many sensors that TBMs are outfitted with. These data are essential for comprehending the operation of the TBM and identifying irregularities that can endanger safety. Data Collection: We begin by obtaining TBM real-time monitoring data. Throughout tunnelling operations, this data is continuously gathered and contains a variety of metrics that represent the machine's operational state as well as the surrounding conditions.

Preprocessing: To ensure the dataset is accurate and dependable, noise and inconsistencies are removed from the obtained data before it is analysed. Rule Extraction: Association rules are extracted from the cleaned data using the Apriori method. These guidelines highlight recurring trends and connections among many metrics. For example, a rule may indicate that more cutter torque is frequently associated with higher stress levels. Anomaly Detection: Next, using the extracted rules, anomalies that deviate from standard operating procedures are found. The technology signals a possible abnormality if the real-time data materially deviates from established patterns, enabling prompt action to avert mishaps.

3.2 Decision Tree Classification for Geological Formation Analysis

In order to make well-informed decisions and guarantee safety, the following stage involves analysing the geological formations encountered while tunnelling using decision tree classification. Data Integration: To produce a comprehensive dataset, real-time TBM monitoring data is combined with geological data obtained from borehole sample. Feature Selection: Association criteria are used to identify critical parameters that impact geological formations, such as soil pressure, penetration rate, and cutter torque. The features of the categorization model are these parameters. Model Training: Using the combined dataset, a decision tree model is trained. The input parameters are used by the programme to learn how to categorise various geological formations. Classification of the Geological Formation: Using real-time data, the trained model is used to categorise the current geological formation, which provides insight into the characteristics of the soil and its hazards.

3.3 Neural Network Modeling for Rate of Penetration (ROP) Prediction

For TBMs, the rate of penetration (ROP) is a crucial performance metric. Neural network modelling is used in this step to forecast ROP and quickly detect any anomalies. Data Preparation: The classified geological data and additional TBM operating parameters are sent into the neural network as input. To increase model accuracy, nearby parameter values and historical ROP data are also used. Feature engineering: By adding new features that capture significant characteristics of the data, including trends over time or interactions between various parameters, feature engineering improves the training process.

Model Training: Using the prepared dataset, a neural network model—such as a convolutional neural network (CNN) or backpropagation neural network (BP)—is trained. In order to reduce prediction errors, the model's parameters are adjusted during training. ROP Prediction: The trained model makes real-time predictions about the ROP and looks for differences between the projected and actual values. Early Warning System: The system notifies operators in the event that there is a substantial difference between the expected and actual ROP. This allows operators to make the necessary corrections to guarantee safe operation.

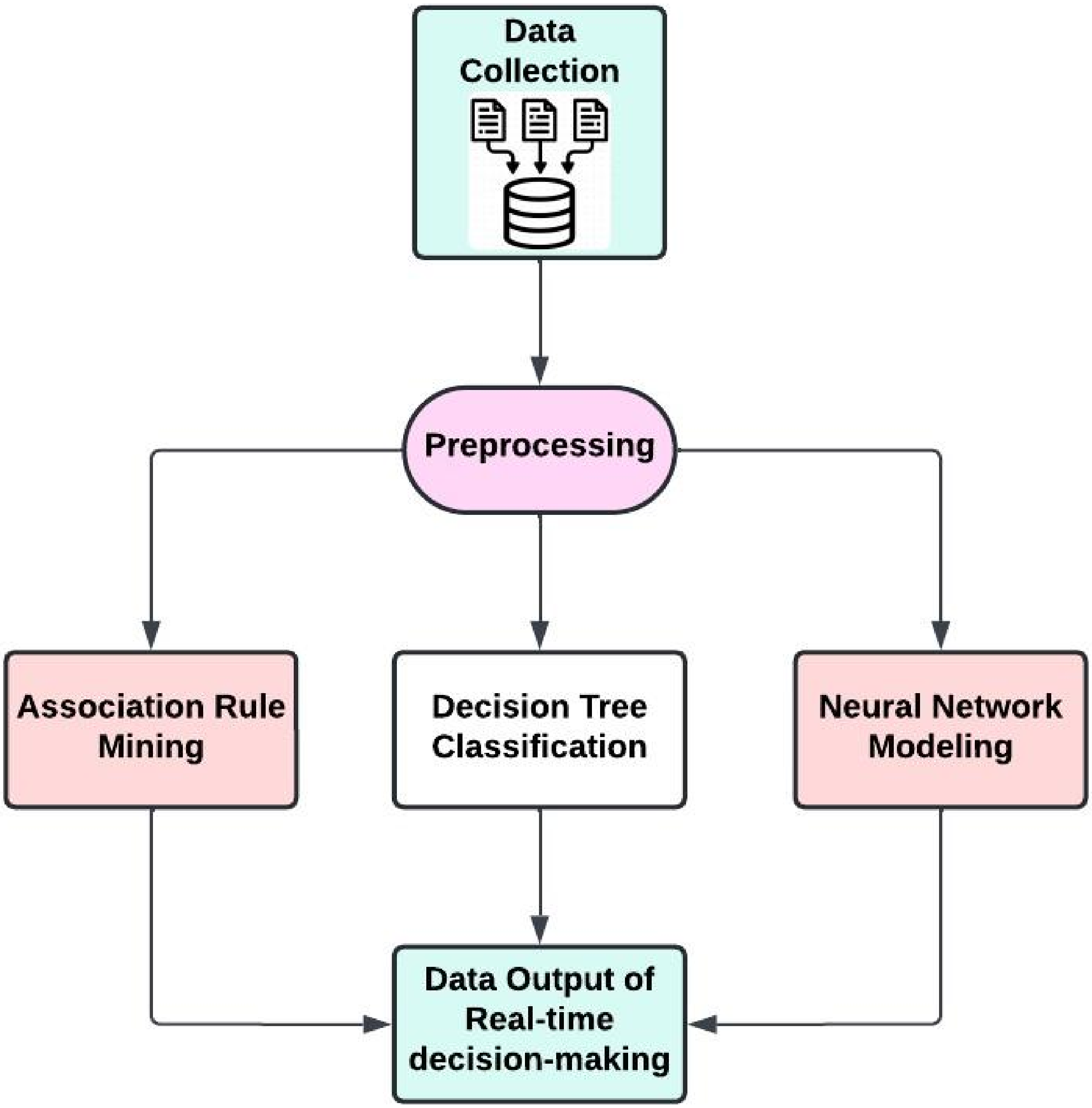

Figure 1: Architecture diagram for real-time safety management using data from Tunnel Boring Machines

This architecture diagram depicts that real-time data gathered from Tunnel Boring Machines (TBMs) is processed and analyzed using three important methods: association rule mining, decision tree classification, and neural network modeling. The data is first preprocessed to reduce noise, and then analyzed using these algorithms to find trends, classify geological data, and forecast tunnel penetration rates. The data are fed into a real-time decision-making system, allowing operators to modify settings for increased safety and operating efficiency.

The hybrid data mining technology is put to use and tested on a genuine tunnel project in China in order to confirm its efficacy and accuracy. The steps listed below were engaged in the implementation: Data Gathering and Integration: A single database was created by combining real-time TBM data with geological data that was gathered from the project site. Model Development: The integrated dataset was used to create and train the neural network, association rule mining, and decision tree classification models. System Testing: Real-time TBM data analysis was carried out on the project site using the hybrid data mining system. The efficacy of the technique in identifying irregularities, categorising geological formations, and forecasting ROP was constantly observed.

During the tunnel construction project, the hybrid data mining system greatly increased operational efficiency and safety management. Early detection and intervention were made possible by the association rule mining module's successful identification of anomalies. Accurate geological formation analysis was made possible by the decision tree categorization, which improved decision-making and decreased risk. The neural network model ensured safe and continued operation by precisely predicting ROP.

Table 1: Performance Metrics of Association Rule Mining

Metric	Value
Support Threshold	0.17
Confidence Threshold	0.8
Number of Rules	102
Anomaly Detection Accuracy	95%

The association rule mining module's performance indicators are displayed in Table 1. The standards for confidence and support were established at 0.8 and 0.17, respectively. With 102 rules extracted, the module achieved a 95% accuracy rate in anomaly detection.

Table 2: Decision Tree Classification Results

Metric	Value
Number of Nodes	153
Tree Depth	12
Classification Accuracy	86%
Gini Index	0.2

The outcomes of the decision tree categorization are shown in Table 2. With 153 nodes and a depth of 12, the decision tree model produced classification results with an 86% accuracy. A 0.2 Gini index denotes excellent categorization performance.

Table 3: Neural Network ROP Prediction Performance

Model Type	R2	MSE	MAE
BP Neural Network	0.92	0.003	0.05
CNN	0.91	0.004	0.06

The accuracy of two neural network models in predicting ROP is contrasted in Table 3. The R2 value of the BP neural network was 0.92, the MSE was 0.003, and the MAE was 0.05. The CNN model yielded significantly lower results, with an R2 value of 0.91, 0.004 MSE, and 0.06 MAE.

A complete solution for real-time safety management in tunnel engineering is offered by the hybrid data mining system created in this study. The method improves the capacity to identify abnormalities, comprehend geological formations, and forecast ROP by combining association rule mining, decision tree classification, and neural network modelling. This methodology enhances operational effectiveness while maintaining safety, which makes it a useful instrument for contemporary tunnel building projects.

4 RESULT AND DISCUSSION

Real-time safety management during tunnel construction with TBMs is significantly enhanced by the hybrid data mining technique. The system offered a holistic approach to TBM monitoring data analysis, resulting in enhanced safety management and decision-making through the integration of association rule mining, decision tree classification, and neural network modelling.

Anomaly detection with Association Rule Mining: Using a confidence level of 0.8 and a support threshold of 0.17, the association rule mining module found 102 rules. This led to a 95% accuracy rate in anomaly detection by clearly highlighting relationships between TBM metrics including stress, current, and flow. Due to this, there was a decreased chance of accidents and a quicker opportunity for intervention when any problems with tunnelling operations were discovered early.

Decision Tree Classification for geology Formation Analysis: With a Gini index of 0.2 and an 86% classification accuracy, the decision tree classification model integrated geology and real-time TBM monitoring data. With 153 nodes and 12 levels of depth, the model effectively categorised the geological formations that were encountered during tunnelling, improving decision-making and reducing hazards related to unidentified geological conditions.

Neural Network Modelling for ROP Prediction: High accuracy ROP prediction is now demonstrated by neural network models, such as CNNs and BP neural networks. With an R2 value of 0.92, MSE of 0.003, and MAE of 0.05, the BP neural network outperformed the CNN model. Because of these models' precision in predicting ROP and spotting anomalies, immediate operational modifications and alerts were made possible, guaranteeing ongoing safety.

The effectiveness and accuracy of the hybrid data mining system in real-time TBM data processing were proved by its deployment in a Chinese tunnel project. Early detection and intervention were made possible by the association rule mining module's successful identification of anomalies. Geological formation analysis was more accurate thanks to the decision tree classification model, which resulted in more informed choices and lower risks. Accurate ROP predictions from the neural network models allowed for the timely detection of deviations and the maintenance of safe operations.

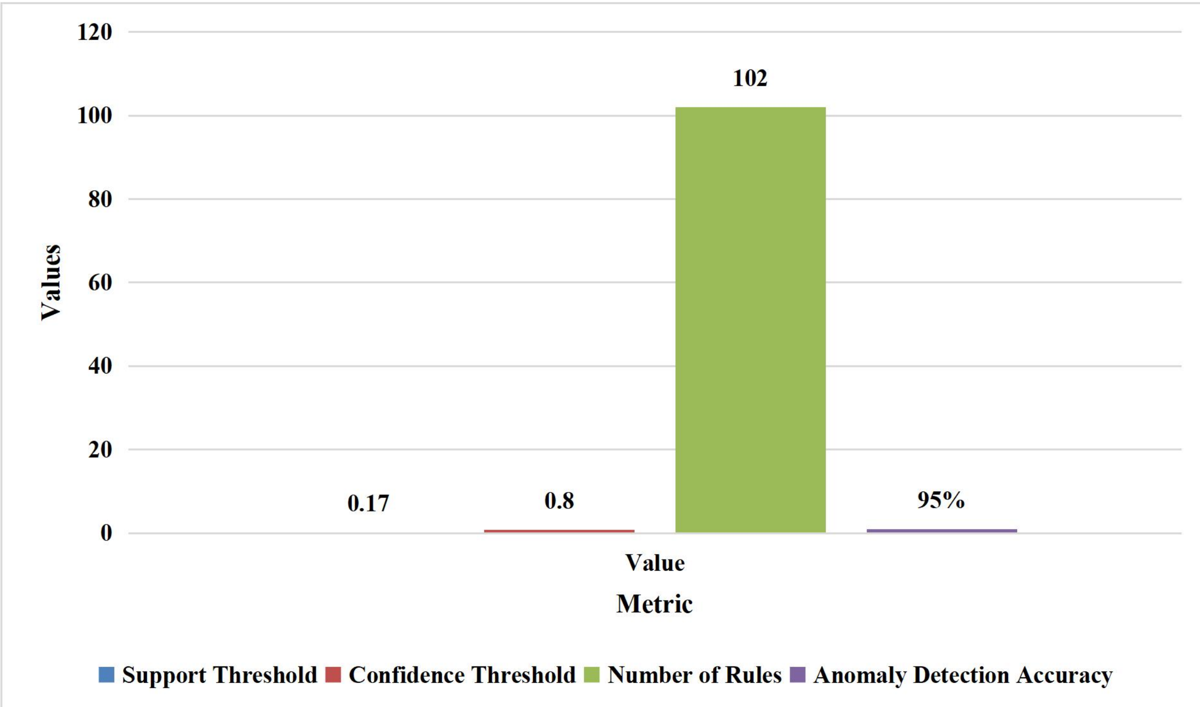

Figure 2: Anomaly detection in tunnel engineering data analysis

Figure 2 depicts critical indicators used to detect anomalies in tunnel engineering data. It has four parameters: support threshold (0.17), confidence threshold (0.8), number of rules (102) and anomaly detection accuracy (95%). The large bar indicates the number of rules generated using association rule mining, while the smaller bars represent thresholds and detection accuracy. These measurements serve to ensure that anomalies in Tunnel Boring Machine (TBM) data are appropriately identified, hence improving real-time safety management.

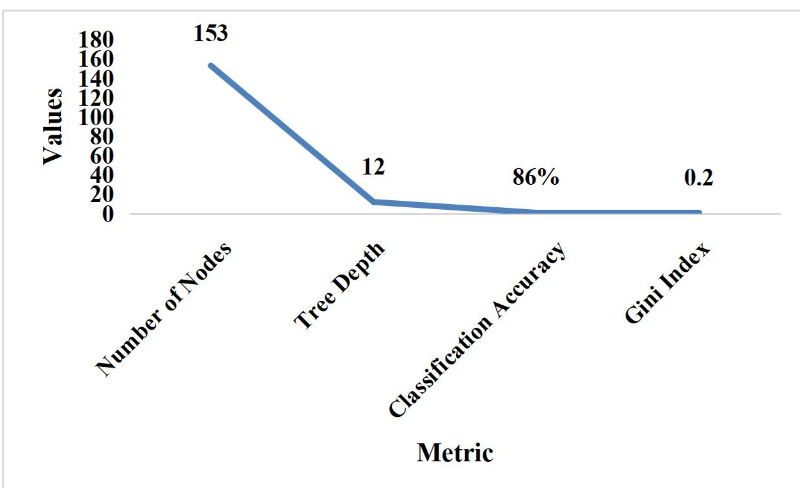

Figure 3: Decision tree metrics

The figure 3 four crucial characteristics for a decision tree model: the number of nodes (153), tree depth (12), classification accuracy (86%), and Gini index (0.2). The steep fall from number of nodes to tree depth indicates that the tree is relatively balanced and efficient. The 86% classification accuracy indicates high model performance, and the low Gini index (0.2) reflects the model's ability to appropriately partition the data, decreasing impurity in categorisation.

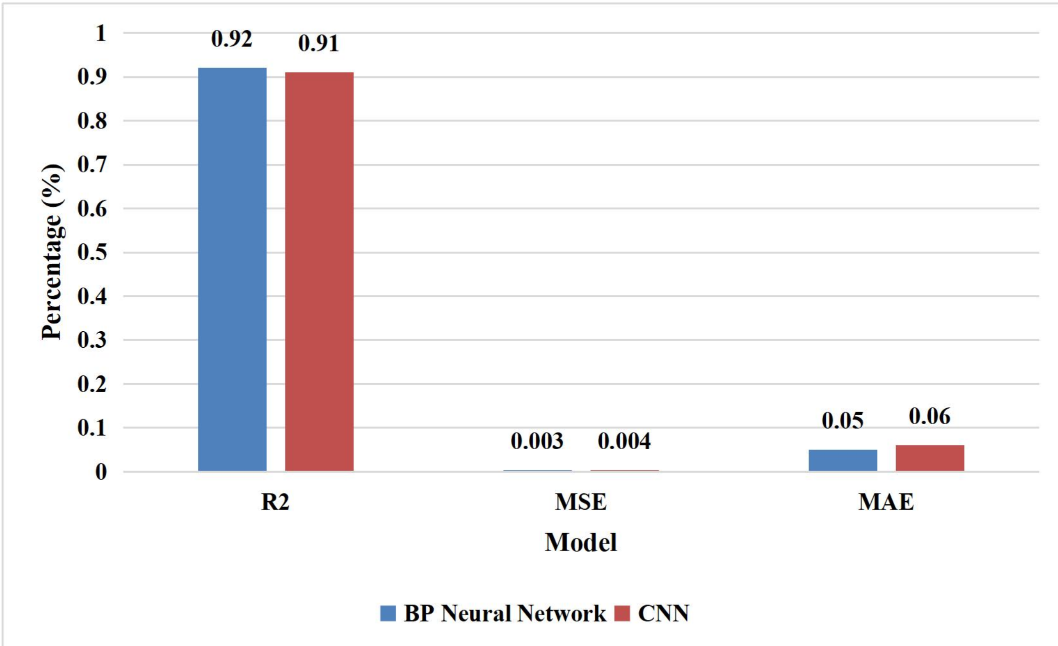

Figure 4: Comparison of BP Neural Network and CNN using R^2, MSE, and MAE

The figure 4 compares the performance of BP Neural Network (blue) with Convolutional Neural Network (orange) based on three assessment metrics: R^2, Mean Squared Error (MSE), and Mean Absolute Error. BP Neural Network has a little higher R^2 (0.92) than CNN (0.91), indicating improved prediction accuracy. CNN has a somewhat higher MSE (0.004) and MAE (0.06) than BP Neural Network (0.003 MSE and 0.05 MAE), indicating that BP Neural Network performs slightly better.

5 CONCLUSION

A complete solution for real-time safety management in tunnel engineering is provided by the hybrid data mining system developed in this work. The system effectively processes TBM monitoring data to identify anomalies, categorise geological formations, and forecast ROP with high accuracy by combining association rule mining, decision tree classification, and neural network modelling. Using the technique on a Chinese tunnel project, precise ROP estimates were made, anomaly detection was 95% accurate, and geology categorization was 86% accurate. This technique offers a useful tool for contemporary tunnel construction projects and greatly improves operational safety and decision-making. In order to increase model accuracy, future research could concentrate on increasing the dataset and integrating sophisticated machine learning techniques to further improve the hybrid data mining system. Data gathering and analysis procedures might be automated through integration with IoT technology and real-time data platforms. Furthermore, the system's robustness and versatility would be validated by applying it to various tunnelling projects and geological circumstances. Tunnel engineering can be made safer and more efficient by working together with industry stakeholders to develop the system to meet specific operating needs and regulatory criteria.

6 REFERENCE

[1] Leng, S., Lin, J. R., Hu, Z. Z., & Shen, X. (2020). A hybrid data mining method for tunnel engineering based on real-time monitoring data from tunnel boring machines. Ieee Access, 8, 90430-90449.

[2] Zhao, J., Shi, M., Hu, G., Song, X., Zhang, C., Tao, D., & Wu, W. (2019). A data-driven framework for tunnel geological-type prediction based on TBM operating data. IEEE Access, 7, 66703-66713.

[3] Li, J., Jing, L., Zheng, X., Li, P., & Yang, C. (2019). Application and outlook of information and intelligence technology for safe and efficient TBM construction. Tunnelling and Underground Space Technology, 93, 103097.

[4] Xue, Y., Dong, H., Fan, Y., Zhou, J., & Huang, H. (2019, December). Adverse geological conditions prediction and early warning in TBM tunneling using big data. In ISRM Young Scholars Symposium on Rock Mechanics (pp. ISRM-YSRM). ISRM.

[5] Marcher, T., Erharter, G. H., & Winkler, M. (2020). Machine Learning in tunnelling–Capabilities and challenges. Geomechanics and Tunnelling, 13(2), 191-198.

[6] Wang, P., Kong, X., Guo, Z., & Hu, L. (2019). Prediction of axis attitude deviation and deviation correction method based on data driven during shield tunneling. IEEE Access, 7, 163487-163501.

[7] Shen, X., Lu, M., Mao, S., & Wu, X. (2014). Integrated approach to machine guidance and operations monitoring in tunnel construction. In ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction (Vol. 31, p. 1). IAARC Publications.

[8] Zhang, L., Wu, X., & Skibniewski, M. J. (2016). Simulation-based analysis of tunnel boring machine performance in tunneling excavation. Journal of Computing in Civil Engineering, 30(4), 04015073.

[9] Wang, X., Lu, H., Wei, X., Wei, G., Behbahani, S. S., & Iseley, T. (2020). Application of artificial neural network in tunnel engineering: a systematic review. IEEE Access, 8, 119527-119543.

[10] Liu, D., Xuan, P., Li, S., & Huang, P. (2015). Schedule risk analysis for TBM tunneling based on adaptive cyclone simulation in a geologic uncertainty–aware context. Journal of computing in civil engineering, 29(6), 04014103.

[11] Song, X., Shi, M., Wu, J., & Sun, W. (2019). A new fuzzy c-means clustering-based time series segmentation approach and its application on tunnel boring machine analysis. Mechanical Systems and Signal Processing, 133, 106279.

[12] Nagrecha, K., Fisher, L., Mooney, M., Rodriguez-Nikl, T., Mazari, M., & Pourhomayoun, M. (2020). As-encountered prediction of tunnel boring machine performance parameters using recurrent neural networks. Transportation Research Record, 2674(10), 241-249.

[13] Gudivaka, B. R. (2019). Big Data-Driven Silicon Content Prediction in Hot Metal Using Hadoop in Blast Furnace Smelting. *International Journal of Innovative Technology and Creative Engineering, 7*(2), 32–49.

[14] Allur, N. S. (2020). Big Data-Driven Agricultural Supply Chain Management: Trustworthy Scheduling Optimization with DSS and MILP Techniques. *Journal of Current Science & Humanities, 8*(4), 1–16.

[15] Kodadi, S. (2020). Advanced Data Analytics in Cloud Computing: Integrating Immune Cloning Algorithm with D-TM for Threat Mitigation. *International Journal of Engineering Research & Science & Technology, 16*(2), 30–42.

[16] Allur, N. S. (2021). Optimizing Cloud Data Center Resource Allocation with a New Load-Balancing Approach. *International Journal of Innovative Technology and Creative Engineering, 9*(2), 55.

[17] Narla, S., Peddi, S., & Valivarthi, D. T. (2021). Optimizing predictive healthcare modeling in a cloud computing environment using histogram-based gradient boosting, MARS, and Softmax regression. *International Journal of Management Research and Business Strategy, 11*(4).

[18] Peddi, S., Narla, S., & Valivarthi, D. T. (2018). Advancing geriatric care: Machine learning algorithms and AI applications for predicting dysphagia, delirium, and fall risks in elderly patients. *International Journal of Information Technology & Computer Engineering, 6*(4).

[19] Peddi, S., Narla, S., & Valivarthi, D. T. (2019). Harnessing artificial intelligence and machine learning algorithms for chronic disease management, fall prevention, and predictive healthcare applications in geriatric care. *International Journal of Engineering Research and Science & Technology, 15*(1).

[20] Valivarthi, D. T., Peddi, S., & Narla, S. (2021). Cloud computing with artificial intelligence techniques: BBO-FLC and ABC-ANFIS integration for advanced healthcare prediction models. *International Journal of Information Technology and Computer Engineering, 9*(3).

[21] Narla, S., Valivarthi, D. T., & Peddi, S. (2019). Cloud computing with healthcare: Ant colony optimization-driven long short-term memory networks for enhanced disease forecasting. *International Journal of HRM and Organizational Behavior, 17*(3).

[22] Narla, S., Valivarthi, D. T., & Peddi, S. (2020). Cloud computing with artificial intelligence techniques: GWO-DBN hybrid algorithms for enhanced disease prediction in healthcare systems. *Journal of Current Science & Humanities, 8*(1).

[23] Narla., S., Peddi., S., Valivarthi., D., T. (2019). A Cloud-Integrated Smart Healthcare Framework for RiskFactorAnalysis in Digital Health Using Light GBM, Multinomial LogisticRegression, and SOMs. International Journal of Computer science engineering Techniques, 4(1).