World Journal of Engineering Research
Print ISSN: 2959-9865
Online ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3073

INTELLIGENT IDENTIFICATION AND DECISION SUPPORT SYSTEM FOR TBM CONSTRUCTION RISK SYNERGISTICALLY DRIVEN BY KNOWLEDGE GRAPHS AND LARGE LANGUAGE MODELS

YongKun Li[1], Yin Bo[2*]
[1]*China Railway Tunnel Bureau Group Co., Ltd., Guangzhou 5111458, Guangdong, China.*
[2]*Changjiang Survey, Planning, Design and Research Co., Ltd., Wuhan 430010, Hubei, China.*
[*]*Corresponding Author: Yin Bo*

Abstract: Tunnel Boring Machines (TBMs) are widely used in subway and tunnel projects due to their efficiency and safety. However, the complexity and high uncertainty of the construction environment lead to lag and limitations in traditional risk identification methods. Addressing the urgent need for intelligent management, this study combines the advantages of Knowledge Graphs (KG) in structured knowledge representation and Large Language Models (LLMs) in natural language understanding to construct a synergistically driven intelligent risk identification and decision support system for TBM construction.First, the study employs ontology modeling to build a risk knowledge graph covering risk factors, risk-causing mechanisms, and consequences, achieving entity-relation extraction through multi-source heterogeneous data mining. Second, the Large Language Model is adapted and fine-tuned by constructing an engineering corpus and injecting domain knowledge. Finally, a synergistic mechanism is designed to realize intelligent risk identification and decision-making.The innovations of this study include: (1) Proposing a "Knowledge Retrieval - Semantic Understanding - Logical Reasoning" three-layer architecture, effectively fusing explicit knowledge with reasoning capabilities; (2) Developing a multi-modal risk feature fusion identification model, improving identification accuracy and real-time performance; and (3) Building an intelligent recommendation engine for risk response plans, supporting multi-scenario simulation. Experimental and engineering application results indicate that the system significantly improves the precision and recall of risk identification, effectively enhancing the risk management level of TBM construction.
Keywords: TBM construction; Risk management; Knowledge graph; Large language model; Intelligent decision support

1 INTRODUCTION

With the continuous expansion of underground space development, Tunnel Boring Machines (TBMs) have become core equipment in tunnel engineering construction due to their efficient and safe construction characteristics. However, facing the increasingly complex and variable underground environments, risk control during TBM construction still encounters severe challenges. The extreme uncertainty of geological conditions—manifested in the unpredictability of disastrous risks such as mud bursts, water inrushes, and rock bursts—superimposed with potential threats of mechanical failures, drastically increases the difficulty of construction safety management. Traditional risk identification methods mostly rely on expert experience rules or statistical analysis based on historical data. The former often exhibits obvious lag and limitations when dealing with dynamic construction scenarios, while the latter struggles to effectively cope with complex, non-linear risk evolution laws. Although some intelligent means have been introduced, existing systems still face key bottlenecks when confronting massive, heterogeneous engineering data, such as difficulties in structural processing, insufficient deep semantic understanding capabilities, and poor real-time performance in decision support.

Meanwhile, breakthrough progress in artificial intelligence technology, particularly the rise of Knowledge Graphs (KGs) and Large Language Models (LLMs), provides a new technical paradigm for solving the aforementioned engineering problems. As a structured knowledge representation method, knowledge graphs can network complex engineering knowledge through ontological modeling, possessing powerful logical reasoning and knowledge retrieval capabilities; meanwhile, LLMs excel in natural language understanding and generation, enabling efficient parsing of unstructured text data such as construction logs and reports. The collaborative drive of both promises to build a three-layer architecture of "knowledge retrieval - semantic understanding - logical reasoning": knowledge graphs provide precise domain background knowledge for LLMs, correcting "hallucination" phenomena; LLMs endow knowledge graphs with stronger generalized understanding and text processing capabilities, thereby achieving deep fusion of unstructured engineering data and structured domain knowledge.

Based on this, this study aims to address key scientific issues existing in current TBM construction risk management, such as the "lack of deep semantic understanding" and "difficulties in multi-source data fusion." This study will explore intelligent risk control solutions suitable for TBM construction environments by constructing an efficient risk knowledge graph and combining it with the collaborative mechanism of large language models. The focus will be on investigating how to utilize the complementary advantages of both to enhance the system's risk identification accuracy and real-time response capability under complex contexts, and to optimize user interaction and decision support functions, ultimately establishing a TBM construction risk intelligent control system with robustness and continuous evolutionary capability.

2 RESEARCH PROGRESS ON TBM CONSTRUCTION RISK INTELLIGENT CONTROL AND KEY TECHNOLOGIES

2.1 Research Progress on TBM Construction Risk Identification

With the continuous development of tunnel engineering technology, the application of TBMs in underground engineering is becoming increasingly widespread. The risks faced during TBM construction possess diversity and complexity; therefore, effective risk identification research is of great significance for ensuring construction safety. Currently, research progress on TBM construction risk identification mainly focuses on the following aspects: First, risk identification methods based on expert experience and rules have been widely used in the past few decades. Such methods typically rely on expert knowledge bases and a series of preset rules to identify potential risks through logical reasoning. However, when dealing with complex and dynamic construction scenarios, these methods often exhibit certain lags and limitations[1]. Second, data-driven risk prediction models based on machine learning have gradually become a research hotspot. These models predict future risks by analyzing massive amounts of historical data to mine the intrinsic connections between risk factors. For example, algorithms such as Support Vector Machines (SVM), Random Forest (RF), and Neural Networks (NN) have achieved certain results in risk identification. However, the adaptability of these models in dynamic scenarios still requires further research. In recent years, with the development of big data and AI technology, researchers have begun to focus on the adaptability analysis of existing methods in dynamic scenarios, such as how to handle noise and outliers in real-time data streams, how to adjust model parameters to adapt to the constantly changing construction environment, and how to improve the generalization ability of models on sparse datasets. Solving these problems is crucial for improving the accuracy and timeliness of risk identification. Furthermore, some studies have attempted to combine knowledge graphs with machine learning models to improve the efficiency and accuracy of risk identification[2]. By constructing a TBM construction risk knowledge graph, a comprehensive understanding of risk factors, hazard mechanisms, and consequences can be achieved, thereby providing more precise input information for machine learning models, as shown in Figure 1. In summary, the progress of TBM construction risk identification research indicates that although existing methods have achieved certain results in theory and practice, challenges remain regarding adaptability and real-time performance in dynamic scenarios. Future research needs to further explore data-driven intelligent methods and combine them with technologies like knowledge graphs to improve the comprehensiveness and accuracy of risk identification.

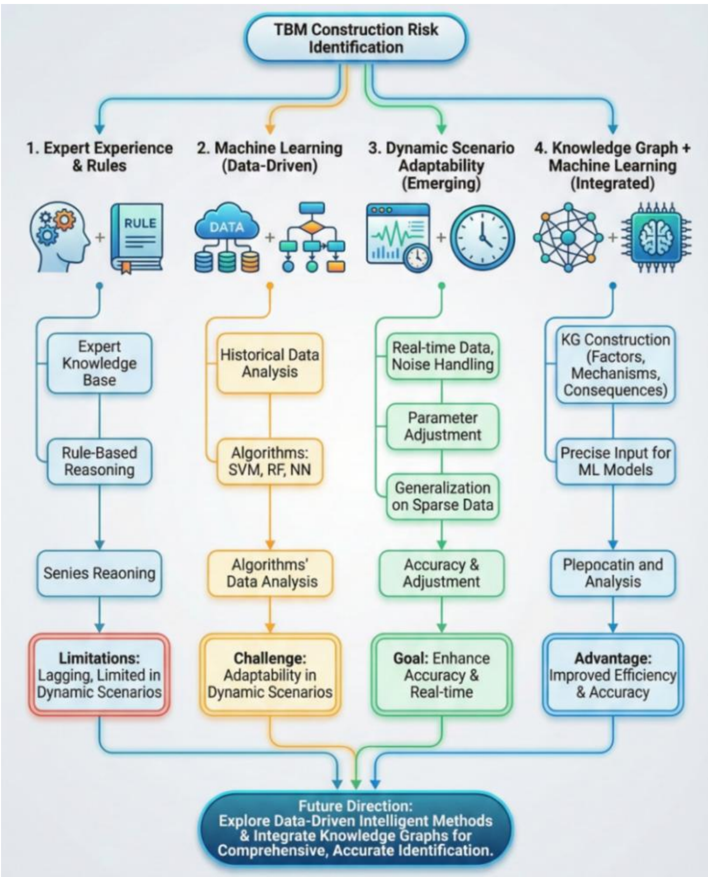

Figure 1 TBM Construction Risk Identification Research Progress

2.2 Overview of Knowledge Graph Construction and Reasoning Technology

As a structured semantic knowledge base, the knowledge graph has seen increasingly widespread application in the field of civil engineering, particularly demonstrating significant advantages in the context of Tunnel Boring Machine (TBM) construction risk management. The construction of a knowledge graph involves multiple technical links, including ontology modeling, entity-relation extraction, and knowledge fusion. Ontology modeling serves as the foundation of knowledge graph construction by defining concepts, attributes, and relations to build a domain knowledge framework. In the field of civil engineering, the TBM construction risk ontology model encompasses multiple dimensions such as risk types, risk factors, hazard mechanisms, and consequences. For instance, a risk ontology concerning tunnel collapse might include geological conditions, construction methods, and environmental factors as risk factors, where the relationships and attributes among them constitute the ontology model. Entity-relation extraction is a key technology in knowledge graph construction, aimed at identifying key entities and their interrelationships from unstructured data[3]. In TBM construction logs and geological reports, the utilization of deep learning technologies, such as Conditional Random Fields (CRF) or Transformer-based models, can effectively identify risk-related entities and relations. Furthermore, combining domain dictionaries with distant supervision techniques can enhance the accuracy and efficiency of relation extraction. Knowledge fusion is the process of handling redundant and contradictory information within the knowledge graph. In TBM construction risk knowledge graphs, information conflicts from different data sources may exist, such as inconsistencies between sensor data and manual records. Through entity alignment and attribute fusion technologies, this information can be integrated to form a consistent knowledge base. Application cases of knowledge graphs in the civil engineering field include, but are not limited to: the construction of risk warning systems that identify potential risks and issue warnings through the analysis of real-time and historical data; the development of decision support systems that utilize knowledge graphs to provide structured knowledge support to assist engineers in decision-making; and construction safety monitoring, which analyzes risk changes during the construction process via knowledge graphs to adjust construction schemes in real time. The application of reasoning technology in knowledge graphs primarily refers to using the structured knowledge within the graph for logical reasoning and prediction[4]. In TBM construction risk management, reasoning technology can be used to predict the probability of risk occurrence, assess the severity of risk consequences, and generate risk response strategies. For example, through path reasoning in the knowledge graph, the interactions between different risk factors can be analyzed to infer potential risk chains. In summary, knowledge graph construction and reasoning technologies play an important role in TBM construction risk management; they not only improve the accuracy of risk identification but also provide a strong knowledge foundation for decision support.

2.3 Application Status of Large Language Models in Engineering Decision-Making

In recent years, Large Language Models (LLMs), as a significant achievement in the field of artificial intelligence, have gradually expanded their application scope from internet content generation to the domain of engineering decision-making. In engineering decision-making, the use of LLMs is mainly reflected in the rapid understanding, processing, and generation of structured knowledge from unstructured data. In the field of civil engineering, the capability boundaries of general large language models have been explored to a certain extent. For example, models can process and analyze text data such as engineering reports and construction logs to quickly identify key information and provide decision support for engineers. Meanwhile, domain-specific large models can further improve their application effects in specific engineering scenarios through fine-tuning and adaptation technologies. Specifically regarding assisted engineering decision-making, the exploratory applications of LLMs are manifested in several aspects. First, LLMs can understand engineering-related professional terminology and complex contexts through natural language processing technologies, providing support for the semantic understanding of engineering problems. Second, LLMs can be combined with knowledge graphs to enhance the semantic understanding and reasoning capabilities of engineering decisions through the structured knowledge of the graph and the generative capability of the LLM. For instance, in TBM construction risk identification, LLMs can assist in analyzing unstructured text in construction logs, extracting risk-related information, and combining it with structured knowledge in the knowledge graph to perform reasoning and prediction of risk factors[5]. Additionally, LLMs can generate solution recommendations for specific engineering problems based on historical data and expert experience. However, despite the progress made in the application of LLMs in engineering decision-making, their application in the engineering field still faces challenges. For example, the generalization ability of models is limited; for rare or complex engineering problems, models may not be able to provide accurate decision support. Furthermore, the real-time performance and accuracy of models also need to be further improved to meet the needs of rapid decision-making at engineering sites[6]. Overall, the application of LLMs in engineering decision-making is still in the primary stage, with huge potential for future development. With the continuous advancement of model technology and the accumulation of engineering data, it is expected that LLMs will play an increasingly important role in engineering decision-making.

2.4 Research Gaps in Collaborative Intelligent Systems in Civil Engineering

As an innovative means of risk management in civil engineering, the research gap in collaborative intelligent systems is mainly reflected in the technical bottlenecks of multi-modal data collaborative processing. Currently, data types in the field of civil engineering are diverse, including structured engineering drawings, unstructured construction logs and text reports, and time-series sensor data. The multi-dimensional characteristics and heterogeneous nature of these data pose challenges to the processing capabilities of collaborative intelligent systems. First, at the data fusion level, how to achieve effective integration and information extraction of data from different sources and formats is a weak link in current research. Although existing studies have effectively organized structured data by constructing knowledge graphs, difficulties remain in the processing of unstructured data, especially in understanding complex semantics and implied information in text. Second, at the data processing level, collaborative intelligent systems need to be able to process and analyze large-scale, dynamically changing data streams. However, existing data processing frameworks often fail to adapt to this dynamism and lack effective strategies in data preprocessing, noise identification, and missing value handling[7]. Furthermore, at the model fusion level, how to effectively combine the structured knowledge of knowledge graphs with the ability of language models to process unstructured text to form complementary advantages is a difficult problem in current research. In addition, the interaction mechanisms and collaborative workflows between models are not yet clear, leading to difficulties in achieving expected results in practical system applications. Finally, at the system evaluation and optimization level, there is currently a lack of a comprehensive evaluation system to measure the performance of collaborative intelligent systems in civil engineering applications, including key indicators such as system real-time performance, accuracy, and robustness[8]. At the same time, research on continuous optimization of systems and knowledge update mechanisms is relatively insufficient. Therefore, addressing the aforementioned research gaps, future research should focus on developing efficient multi-modal data processing frameworks, optimizing the collaborative mechanisms of knowledge graphs and language models, and establishing sound system evaluation and optimization processes to promote the widespread application of collaborative intelligent systems in the field of civil engineering.

2.5 Limitations of Existing Research and Breakdown Directions of This Paper

Although existing research on TBM construction risk identification has made certain progress, there are still numerous limitations. First, the problems of data silos and knowledge fragmentation are prominent, leading to incomplete information acquisition during the risk identification process, which affects the accuracy of risk control. Second, traditional models often exhibit insufficient model interpretability and decision reliability when dealing with complex non-linear relationships, making it difficult to meet actual engineering requirements. This paper aims to break through the limitations of existing research and proposes the following innovative paths: first, constructing a comprehensive TBM construction risk knowledge graph to achieve deep fusion of unstructured engineering data and structured knowledge; second, adopting large language models to enhance the system's deep semantic understanding capability of complex engineering texts; third, designing risk identification and decision-making algorithms driven by the collaboration of knowledge graphs and LLMs to improve the accuracy and reliability of decisions[9]. Specifically, this paper will focus on solving the following key problems: how to construct a knowledge graph suitable for TBM construction risk control to achieve comprehensive data fusion and deep mining; how to fine-tune and optimize large language models to improve their semantic understanding and generation capabilities in the engineering domain; and how to design effective collaborative mechanisms so that knowledge graphs and LLMs can exert their maximum efficacy in risk identification and decision-making. Through these innovative paths, this study is expected to provide a more intelligent solution for TBM construction risk control.

3 OVERALL SYSTEM ARCHITECTURE DESIGN

3.1 System Design Goals and Functional Requirements

The full-process TBM construction risk intelligent control system aims to achieve real-time monitoring, intelligent identification, and decision support for TBM construction risks by integrating knowledge graphs and large language models[10]. The design goals and functional requirements of the system are as follows: First, the system design goals include improving the real-time performance and accuracy of TBM construction risk management to achieve early warning and effective control of risks; building a user-friendly interactive interface to enhance user decision-making experience and efficiency; and ensuring the reliability and stability of the system to adapt to complex construction environments and diverse user needs. Specifically, the functional requirements of the system are divided into the following aspects: 1. Real-time and Accuracy Performance Indicators: The system needs to possess high-speed data processing capabilities to meet real-time requirements; meanwhile, through deep learning and knowledge reasoning, it ensures the accuracy of risk identification and prediction. 2. User Interaction Function: The system should provide an intuitive and easy-to-operate user interface, including a risk situational awareness dashboard, visualized browsing of knowledge graphs, and a natural language Q&A system, to support users in efficiently acquiring information, submitting data, and receiving decision recommendations. 3. Decision Support Function: The system needs to integrate multiple information resources such as historical data mining, real-time data analysis, and expert knowledge bases, and provide services such as intelligent recommendation of risk response plans, multi-scenario simulation, and contingency plan generation through intelligent algorithms. 4. System Scalability and Compatibility: The system design should adopt a modular plugin design concept to support compatibility with different types of TBM equipment; at the same time, the system should have cloud deployment and localized adaptation capabilities to meet the application needs of different scenarios. 5. System Self-learning and Optimization Mechanism: By collecting user feedback data, the system can achieve self-optimization, continuously improving the accuracy and reliability of decision recommendations, and forming a benign feedback loop. Through the realization of the above design goals and functional requirements, the full-process TBM construction risk intelligent control system will provide strong support for risk management, reduce the risk of safety accidents, and improve construction management efficiency and quality.

3.2 Collaborative Mechanism Framework of Knowledge Graph and LLM

The collaborative mechanism framework of Knowledge Graphs and Large Language Models (LLMs) aims to build an efficient and intelligent risk identification and decision support system. The three-layer architecture proposed in this study, namely "Knowledge Retrieval - Semantic Understanding - Logical Reasoning," provides a complete process for the system from data input to decision output, as shown in Figure 2. First, the knowledge retrieval layer is responsible for extracting structured knowledge related to the current construction environment from the knowledge graph. This process includes queries to the ontology model and efficient retrieval of entities, relations, and attributes. The ontology modeling in the knowledge graph not only covers the conceptual classification system of TBM construction risks but also defines the hierarchical relationships among risk factors, hazard mechanisms, and consequences. Second, the semantic understanding layer performs deep parsing of unstructured engineering text data through the LLM to achieve accurate understanding and expression of risk information. In this layer, the interface design and protocol specification for dual-model interaction are key, ensuring information flow and collaborative work between the knowledge graph and the LLM[11]. The injection of domain knowledge and Prompt engineering optimization further enhance the LLM's ability to understand professional terminology and complex contexts. Finally, the logical reasoning layer utilizes the structured knowledge of the knowledge graph and the semantic understanding capability of the LLM to perform logical reasoning and decision generation. The closed-loop feedback mechanism under collaborative drive not only realizes real-time updates of risk identification results but also continuously optimizes the decision process through model self-learning. This mechanism can effectively handle dynamically changing construction environments and adjust risk response strategies in a timely manner. In the practical application of the framework, the collaborative effect of the knowledge graph and LLM is demonstrated in multiple aspects. For example, the knowledge graph provides structured knowledge support for the LLM, enhancing its reasoning ability when facing complex engineering problems; meanwhile, the semantic understanding capability of the LLM contributes to the dynamic update and quality assessment of the knowledge graph, improving the accuracy and coverage of the knowledge graph. Through this collaborative mechanism framework, the system can achieve real-time monitoring, accurate identification, and effective decision-making for TBM construction risks, providing a new methodology and practical path for intelligent control of construction risks.

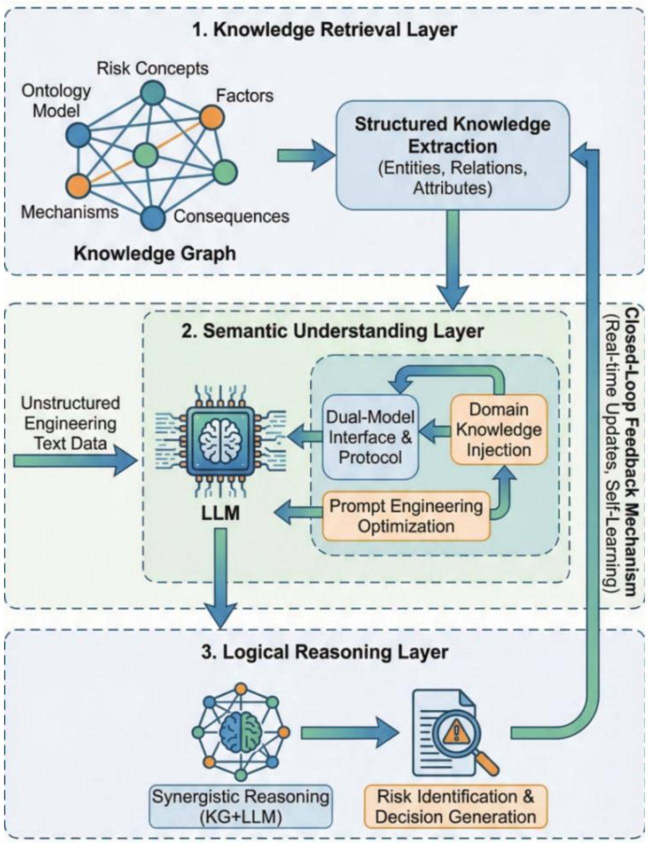

Figure 2 Knowledge Graph & LLM Collaborative Mechanism Framework for TBM Risk Identification

3.3 Data Flow and Module Interaction Logic

Multi-source heterogeneous data undergoes a rigorous input and flow process within the system. First, the system integrates information from multiple channels, including geological investigation reports, construction logs, sensor time-series data, and text data. Before entering the system, these data must undergo structured processing and cleaning alignment to ensure data quality and consistency; data noise identification and missing value imputation strategies are crucial in this process, ensuring the accuracy of subsequent processing. The calling logic between the Knowledge Graph module and the Large Language Model (LLM) module constitutes the core interaction mechanism of the system. The Knowledge Graph module is responsible for extracting key information such as risk factors, hazard mechanisms, and consequences from structured data, and establishing the relational attributes and constraint rules of the ontology model. Meanwhile, the LLM module performs deep semantic understanding on unstructured text data to extract text features related to risks[12]. In the decision result output and visualization process, the system first conducts logical reasoning through the Knowledge Graph module to identify potential risk points and transmits this information to the LLM module. Utilizing its powerful language generation capability, the LLM module translates these logical reasoning results into easy-to-understand natural language descriptions, which are then combined with data visualization technologies to provide users with intuitive risk assessment reports. Furthermore, the system designs a closed-loop feedback mechanism to ensure that the interaction between the Knowledge Graph and the LLM module is dynamic and iterative. This mechanism allows the system to continuously optimize models based on the latest data and feedback, thereby improving the accuracy and efficiency of decision-making. Through this logic design of data flow and module interaction, the system achieves full-process intelligent control of TBM construction risks, meeting performance indicators for real-time capabilities and accuracy, and providing effective decision support functions for users.

3.4 System Scalability and Compatibility Design

System scalability and compatibility design are key factors guaranteeing stable operation in diverse environments. In this study, the system adopts a modular plugin design concept to adapt to different TBM equipment and construction environments. Modular design allows the system to rapidly integrate new functional modules according to specific needs without large-scale refactoring of the existing system, significantly improving system flexibility and scalability. regarding compatibility, the system has designed a series of interfaces to be compatible with data output formats of different types of TBM equipment. These interfaces follow open standards and general protocols, ensuring the system can seamlessly dock with mainstream TBM equipment in the market[13]. Additionally, the system supports adaptive conversion of data formats, capable of handling data from different sources and structures, thereby enhancing data compatibility. The cloud deployment scheme aims to achieve remote access and large-scale data processing for the system. Through cloud deployment, the system can provide flexible scaling capabilities to cope with construction projects of different scales; meanwhile, cloud deployment supports multi-user concurrent access, ensuring efficient system response. The localization adaptation scheme takes into account the potential network restrictions and data privacy issues at construction sites. The system can utilize edge computing technology to migrate partial calculation tasks to the local end, reducing dependency on cloud services. This design not only ensures system real-time performance but also enhances data security. The system's scalability and compatibility design enable it to adapt to the constantly changing technical environment and construction needs, while also facilitating future functional upgrades and technical iterations. In this way, the system can maximize user satisfaction while guaranteeing performance.

4 CONSTRUCTION OF A KNOWLEDGE GRAPH FOR TBM CONSTRUCTION RISKS

4.1 Risk Ontology Modeling and Semantic Definition

In the process of constructing the TBM construction risk knowledge graph, risk ontology modeling and semantic definition are critical steps[14]. First, by constructing a TBM construction risk concept classification system, risk events are classified according to type, cause, and impact, forming a hierarchical conceptual structure. On this basis, hierarchical definitions are applied to risk factors, hazard mechanisms, and consequences to clarify the relationships and attributes among various concepts. Furthermore, the design of relational attributes and constraint rules for the ontology model is an important link in ensuring the quality and usability of the knowledge graph. Relational attributes define the associations between concepts, such as causal relationships and hierarchical relationships. Constraint rules restrict concepts and relationships to ensure the correctness and consistency of knowledge; for example, rules can be set to restrict a certain risk factor to be associated only with specific types of hazard mechanisms. During the ontology modeling process, it is also necessary to consider the attribute definitions of risk factors, such as risk level, occurrence probability, and impact scope. These attributes facilitate quantitative analysis and assessment of risks. In addition, to improve the usability of the knowledge graph, semantic relationships such as synonyms, hypernyms, and hyponyms for related concepts should also be defined. Through the aforementioned ontology modeling and semantic definition, a structured and hierarchical knowledge system can be built for the TBM construction risk knowledge graph. This system helps achieve effective organization and retrieval of risk knowledge, supporting subsequent risk identification, prediction, and decision-making. Simultaneously, the ontology model provides a foundation for the sharing and reuse of risk knowledge, promoting the dissemination and accumulation of knowledge within the domain.

4.2 Multi-source Heterogeneous Data Acquisition and Preprocessing

In the process of constructing the TBM construction risk knowledge graph, the acquisition and preprocessing of multi-source heterogeneous data are critical steps. Multi-source data includes geological investigation reports, construction logs, sensor time-series data, and related text materials, which exhibit significant differences in data type, format, and structure, posing challenges for data integration and processing. First, regarding the structured processing of geological investigation reports and construction logs, this study employs automated information extraction and natural language processing technologies to convert unstructured text data into structured data formats. By designing an ontology model suitable for the engineering domain, key information in the text, such as risk events, risk factors, and construction parameters, can be effectively identified and extracted. Secondly, the cleaning and alignment of sensor time-series data and text data are important links in preprocessing. Since sensor data may be affected by environmental noise and equipment malfunctions, de-noising processing is required. This study adopts time-series analysis and wavelet transform methods to de-noise sensor data, while utilizing data synchronization technology to ensure the correspondence between text data and time-series data in the time dimension. Data noise identification is a difficult point in data preprocessing. This study establishes an anomaly detection model to monitor data in real-time to identify and exclude outliers. Furthermore, addressing the data missingness problem, multiple strategies are adopted for imputation, including mean imputation, interpolation imputation, and model-prediction-based imputation methods, to reduce the impact of missing data on knowledge graph construction. When processing multi-source heterogeneous data, this study also focuses on data consistency and accuracy. By designing data quality assessment indicators, the completeness, accuracy, and consistency of data are quantitatively evaluated to ensure that the quality of preprocessed data meets the requirements for knowledge graph construction. In summary, the acquisition and preprocessing of multi-source heterogeneous data provide a high-quality data foundation for the construction of the TBM construction risk knowledge graph, as shown in Figure 3. This process involves not only data format conversion and cleaning but also multiple links such as noise identification, missing value imputation, and data quality assessment, thereby ensuring the reliability and effectiveness of the knowledge graph.

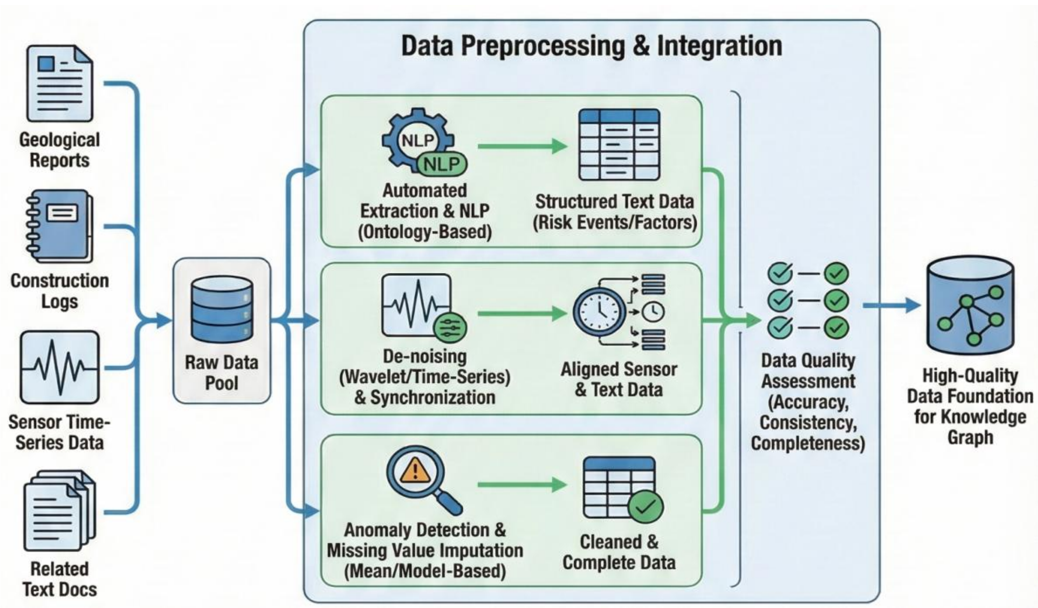

Figure 3 TBM Construction Risk Knowledge Graph: Multi-Source Heterogeneous Data Collection & Preprocessing Workflow

4.3 Entity Extraction and Relationship Mining Methods

In the process of constructing the TBM construction risk knowledge graph, entity extraction and relationship mining are crucial steps. Entity extraction involves identifying key concepts from unstructured text, such as risk events, hazard factors, and construction phases, while relationship mining aims to discover the semantic associations between these entities. This study employs a Named Entity Recognition (NER) model based on deep learning for entity extraction. Based on pre-trained neural networks, this model improves the recognition accuracy of entities related to TBM construction risks by incorporating domain-specific vocabularies and syntactic rules. Furthermore, the model achieves joint training of entity recognition and entity classification through a multi-task learning framework, thereby enhancing overall recognition performance. regarding relationship mining, this study designs a relationship extraction algorithm oriented towards complex contexts. This algorithm first utilizes dependency parsing to identify grammatical dependencies between entities and then combines prior knowledge from the knowledge graph to predict and verify potential semantic relationships. To address the issue of insufficient annotated data in relationship extraction, this study adopts distant supervision technology, which automatically generates pseudo-labeled data from large-scale unannotated texts, expanding the training set size and improving the coverage and accuracy of relationship extraction. Additionally, this study constructs a domain dictionary to enhance the accuracy of entity extraction and relationship mining. This dictionary contains professional terminology and related vocabulary in the field of TBM construction risk, effectively improving model performance by introducing dictionary information during the model training and prediction stages. In the process of entity extraction and relationship mining, this study also introduces a manual review mechanism to resolve potential errors in the model. Through manual review, biases in model predictions can be discovered and corrected in a timely manner, ensuring the quality of knowledge graph construction. In summary, by combining deep learning models, distant supervision technology, domain dictionaries, and manual review mechanisms, this study achieves effective extraction of entities and deep mining of relationships in TBM construction risk texts, providing a solid foundation for knowledge graph construction.

4.4 Graph Dynamic Update and Quality Assessment Mechanism

As an important carrier of structured knowledge, the dynamic update and quality assessment of the knowledge graph are essential for maintaining the timeliness and accuracy of knowledge. In the construction process of the TBM construction risk knowledge graph, the design of the incremental knowledge update algorithm aims to ensure the timely integration of new knowledge while handling conflict and consistency detection between knowledge. This mechanism involves processing newly collected data, including entity recognition, relationship extraction, and attribute extraction, and subsequently integrating the extracted knowledge into the existing graph in an appropriate manner. In the incremental update process, the algorithm needs to prioritize consistency and redundancy detection of knowledge. By establishing a set of strict quality control standards, such as using graph algorithms to detect cycle structures in the knowledge graph, potential redundant relationships can be identified and eliminated. Furthermore, consistency detection rules are introduced to ensure that newly added knowledge does not conflict with information in the existing knowledge base. Quantitative assessment indicators for graph coverage and accuracy are important dimensions for measuring graph quality. The coverage indicator is used to measure the extent of knowledge coverage within a specific domain, typically assessed by calculating the ratio of annotated entities and relationships to the total potential entities and relationships. The accuracy indicator focuses on the correctness of knowledge in the graph, verified through manual annotation or comparison with other authoritative data sources. To improve the objectivity and automation level of assessment, an automated assessment framework can be designed. This framework periodically extracts sample data from external authoritative data sources and conducts comparative analysis with data in the knowledge graph to generate detailed reports on graph quality. Additionally, machine learning algorithms can be utilized to predict knowledge in the graph and compare it with actual results to assess the prediction accuracy of the graph. Assessment results not only help in understanding the quality of the graph but also provide guidance for the optimization of the knowledge graph. Based on assessment results, knowledge extraction algorithms can be adjusted, knowledge fusion strategies optimized, or existing knowledge corrected and supplemented. In this way, the knowledge graph can be continuously iteratively improved, providing more accurate and comprehensive support for TBM construction risk control.

5 LARGE LANGUAGE MODEL ADAPTATION AND FINE-TUNING STRATEGIES

5.1 Construction of Engineering Domain Corpus

As a fundamental resource for natural language processing, the corpus is crucial for the adaptation and fine-tuning of large language models in the engineering domain. The engineering domain corpus constructed in this study aims to provide rich, professional data support with practical application backgrounds for the model. The construction of the corpus includes the establishment of a TBM professional terminology database and a construction case database, as well as the cleaning and annotation of high-quality instruction fine-tuning data. The construction of the TBM professional terminology database is based on in-depth analysis of literature in the tunnel engineering field, collecting and organizing professional terms related to TBM construction, including equipment components, construction processes, and geological conditions. The construction case database gathers typical TBM construction cases from home and abroad, covering application scenarios under different geological conditions and different types of TBM equipment, as well as corresponding risk events and response measures. The cleaning and annotation of high-quality instruction fine-tuning data constitute a key link in corpus construction[15]. The cleaning process involves removing duplicate data, correcting erroneous data, and filtering irrelevant information. The annotation process employs professional knowledge and domain expert experience to precisely classify and annotate data, ensuring data quality meets the requirements of model training. Version management of corpus data is significant for tracking data changes and ensuring data consistency. This study adopts a version control system to manage different versions of the corpus, ensuring detailed records for each update to facilitate backtracking and problem localization. regarding quality control, a data quality assessment system is established to periodically evaluate the quality of the corpus, ensuring data accuracy and reliability. Furthermore, the corpus construction process also emphasizes data diversity and balance to avoid model overfitting on specific types of data. Through the aforementioned construction process, this study provides a high-quality data foundation for the adaptation and fine-tuning of large language models in the engineering domain.

5.2 Domain Knowledge Injection and Prompt Engineering Optimization

Domain knowledge injection is a key technology for enhancing the performance of large language models in specific application scenarios. This study adopts a knowledge injection method based on Retrieval-Augmented Generation

(RAG), combining the pre-constructed TBM construction risk knowledge graph with the language model to improve the model's ability to understand and generate professional domain texts. regarding Prompt engineering optimization, this study designs a dynamic Prompt generation strategy. This strategy dynamically adjusts Prompt content based on the contextual information of the input text, optimizing the context window to better guide the model to focus on task-related information. Additionally, Chain-of-Thought (CoT) prompt design facilitates the model in performing logical reasoning and producing semantically coherent outputs during generation. specifically, this study first extracts key risk factors and hazard mechanisms by analyzing entity relationships and attributes in the TBM construction risk knowledge graph. This information is used to construct Prompt templates, which, combined with general knowledge learned during the model's pre-training phase, enhance the model's understanding capabilities in the professional domain. During the dynamic Prompt generation process, the system adjusts the contextual information of the Prompt in real-time according to the user's input text content to include knowledge most relevant to the current task. This strategy helps the model quickly locate key information and reduce interference from irrelevant information when handling complex engineering problems. Furthermore, this study also explores the impact of context window size on model performance. Experiments reveal that an appropriate context window size can significantly improve risk identification accuracy, while windows that are too large or too small lead to performance degradation. In summary, domain knowledge injection and Prompt engineering optimization effectively improve the performance of large language models on TBM construction risk identification tasks, laying a foundation for the model's practical application.

5.3 Model Fine-tuning Methods and Parameter Configuration

In the model fine-tuning stage, selecting an appropriate pre-trained model as a base is critical. This paper selects and evaluates various pre-trained language models based on the characteristics of the TBM construction field. On this basis, efficient fine-tuning techniques are explored, and model hyperparameters are optimized. First, this paper evaluates the adaptability of mainstream pre-trained models such as BERT, RoBERTa, and GPT-3 in the TBM construction field. Comparative experiments reveal that the RoBERTa model possesses better semantic understanding and generation capabilities when processing text data in the construction field. Therefore, this paper selects RoBERTa as the base model for subsequent fine-tuning. Second, this paper adopts efficient fine-tuning techniques such as LoRA (Low-Rank Adaptation) and P-Tuning to reduce model training costs and accelerate fine-tuning speed. LoRA adapts model parameters by introducing low-rank matrices, while P-Tuning achieves fine-tuning by adjusting the output distribution of the model's intermediate layers[16]. Both methods can improve fine-tuning efficiency without significantly affecting model performance. regarding hyperparameter optimization, this paper employs Grid Search and Bayesian Optimization methods. Comparative experiments show that Bayesian Optimization has higher efficiency and better optimization effects during the parameter space search process. Specifically, this paper optimizes the following hyperparameters: 1. Learning rate: Choosing an appropriate learning rate is crucial for model training. An excessively high learning rate may cause the model to fail to converge, while an excessively low learning rate may lead to an overly long training process. 2. Training batch size: An appropriate batch size can balance the convergence speed of model training and memory consumption. 3. Training epochs: Increasing training epochs can improve model performance but also increases computational costs. Additionally, this paper adjusts training strategies, such as adopting dynamic learning rate adjustment and gradient accumulation techniques, to improve model performance and stability during the fine-tuning process. In summary, through pre-trained model selection, efficient fine-tuning techniques, and adjustments to hyperparameter optimization and training strategies, this paper successfully improves the adaptability and performance of large language models in the TBM construction field, providing an effective model foundation for subsequent risk identification and decision support.

5.4 Verification of Risk Semantic Understanding and Generation Capabilities

In the verification of risk semantic understanding and generation capabilities, this study mainly focuses on the accuracy of the model in understanding complex engineering texts and the professionalism and fluency in generating risk descriptions. First, by analyzing a series of TBM construction logs, geological reports, and expert interview records, a text dataset containing real engineering scenarios was constructed. This dataset covers various risk types, such as collapses, water inrushes, and mechanical failures, as well as corresponding risk descriptions and response measures. To assess the model's risk semantic understanding capability, this study designed a series of qualitative and quantitative tests. Qualitative testing involves expert review, where domain experts are invited to score the risk descriptions generated by the model to judge whether they conform to engineering reality and professional standards. Quantitative testing uses accuracy, recall, and F1 score as evaluation indicators to measure the model's performance in identifying and generating risk descriptions[17]. In the tests, the model demonstrated high professionalism and fluency in risk description generation, but hallucination phenomena still existed in certain complex contexts, where the model might generate descriptions that seem reasonable but are actually inconsistent with the context. To suppress this phenomenon, this study adopted the following strategies: 1. Introduce domain knowledge-enhanced pre-trained models to improve the model's understanding of professional terminology and complex relationships by fusing structured knowledge from the knowledge graph with the capabilities of the language model. 2. Optimize Prompt design by introducing Chain-of-Thought prompts and dynamic Prompt generation strategies to guide the model to pay more attention to contextual information when generating descriptions. 3. Apply model hallucination detection technology to identify and correct potential hallucination content by analyzing the correlation between generated risk descriptions and facts in the knowledge graph. Test results indicate that after optimization with the above strategies, the hallucination phenomenon in the model's risk description generation was effectively suppressed, and accuracy, recall, and F1 score all improved. Especially when processing engineering texts containing rich professional terminology and complex relationships, the model's performance significantly improved, showing good adaptability and robustness. In summary, through meticulous model assessment and optimization, this study verified the effectiveness of the model in terms of risk semantic understanding and generation capabilities, providing reliable technical support for intelligent control of TBM construction risks. However, the model's performance in extremely rare risk scenarios remains to be further studied and improved.

6 COLLABORATIVE DRIVING MECHANISM AND RISK IDENTIFICATION ALGORITHMS

6.1 Knowledge Graph-Guided LLM Reasoning Enhancement Strategy

The Knowledge Graph-guided Large Language Model (LLM) reasoning enhancement strategy aims to improve the LLM's understanding and reasoning capabilities for complex engineering problems through structured knowledge. The core of this strategy lies in mapping structured knowledge from the knowledge graph to the LLM's process of handling unstructured text, thereby enhancing the model's reasoning capability in specific domains. First, the mapping mechanism from structured knowledge to unstructured text involves combining elements such as entities, relationships, and attributes from the knowledge graph with the LLM's input text. By embedding knowledge from the graph into the text in the form of semantic role labeling, rich background knowledge is provided to the LLM, facilitating its understanding and reasoning of implied information in the text. Second, the reasoning correction method based on graph path constraints is one of the key technologies of this strategy. This method utilizes path information in the knowledge graph to constrain and guide the LLM's reasoning process, reducing potential biases and errors during reasoning. By analyzing relationship paths between entities in the graph, the LLM can more accurately identify and infer implied relationships and logic in the text. Furthermore, knowledge tracing technology to improve model interpretability is an important component of the strategy. By performing knowledge tracing on the LLM's reasoning results, the specific knowledge sources supporting or refuting a certain conclusion can be tracked, thereby improving the interpretability of model decisions. This process not only helps enhance user understanding and trust in model decisions but also provides effective means for model debugging and optimization. In summary, the Knowledge Graph-guided LLM reasoning enhancement strategy provides new methods and ideas for intelligent reasoning in the engineering field by fusing structured knowledge with the LLM's text processing capability. The implementation of this strategy is expected to significantly improve the performance and reliability of LLMs when dealing with complex engineering problems.

6.2 LLM-Assisted Knowledge Graph Completion and Correction Mechanism

Knowledge graphs possess significant advantages in structured knowledge representation and reasoning; however, practically constructed knowledge graphs often suffer from incompleteness and errors, which limit their application effectiveness. To this end, this study proposes an LLM-assisted knowledge graph completion and correction mechanism. First, utilizing the text generation capability of the LLM, potential knowledge can be extracted from unstructured text to complete missing relationships in the knowledge graph. Specifically, by designing appropriate Prompts, the LLM is guided to generate sentences describing entity relationships, from which triples are then extracted using entity recognition and relationship extraction technologies and added to the knowledge graph. Second, this study adopts a method based on semantic logic for anomaly knowledge identification and correction. By representing relationships in the knowledge graph as logical expressions and combining them with the LLM's reasoning capability, logical contradictions and anomalous knowledge can be detected. For detected anomalous knowledge, the system automatically performs corrections to ensure the accuracy and consistency of the knowledge graph. Furthermore, this study designs a bidirectional iterative optimization process for the graph and the model. On the one hand, LLM-assisted knowledge graph completion and correction can improve the quality of the knowledge graph, thereby enhancing the performance of the entire system; on the other hand, optimization of the knowledge graph can conversely guide the fine-tuning and optimization of the LLM, forming a virtuous cycle. In practical application, this study selected a knowledge graph in the tunnel engineering field for testing. Experimental results show that the proposed LLM-assisted knowledge graph completion and correction mechanism can effectively improve the completeness and accuracy of the knowledge graph, thereby enhancing the accuracy of risk identification and decision-making.

6.3 Multi-modal Risk Feature Fusion Recognition Model

The multi-modal risk feature fusion recognition model aims to build a more comprehensive risk assessment framework by integrating text semantic features and numerical time-series features. Text semantic features can capture potential risk information in unstructured data such as construction logs and reports, while numerical time-series features can reflect dynamic trends in sensor data. The combination of both not only enhances the model's recognition capability but also improves its adaptability in complex scenarios. In the fusion architecture, text data undergoes preprocessing and entity recognition, followed by the extraction of key semantic information such as risk events and hazard factors via natural language processing technologies. Meanwhile, numerical time-series data undergoes feature engineering, such as time window extraction and statistical feature calculation, converting into quantifiable risk indicators. The application of a cross-modal attention mechanism enables the model to effectively associate risk features across different modalities, highlighting key information through weight allocation, and subsequently improving risk identification accuracy. Empirical research indicates that the fusion model demonstrates significant advantages when processing multi-source heterogeneous data8. For example, in a certain TBM construction project, the fusion model successfully identified potential risks difficult for traditional methods to discover by analyzing geological reports and sensor data. Additionally, the robustness of the fusion model against data sparsity and noise was also verified; even under conditions of incomplete or interfered information, the model maintained high recognition accuracy. Analysis of advantages in multi-source data complementarity shows that the fusion model can synthesize information from different data sources to form a more comprehensive risk profile. Text data provides qualitative descriptions of risks, while numerical data supplements quantitative dimensions; the combination of both makes risk identification more precise. Furthermore, the fusion model can establish connections between different types of risks, providing deeper insights for risk management and decision-making.

6.4 Real-time Risk Scoring and Prioritization Algorithm

In the design of the real-time risk scoring and prioritization algorithm, this study proposes a comprehensive evaluation model based on multi-attribute decision-making. This model comprehensively considers multiple dimensions such as potential impact, probability of occurrence, and urgency of risks to achieve dynamic quantitative assessment and prioritization of TBM construction risks. First, the model quantifies various risk factors by constructing a risk factor indicator system. Specific methods include using expert scoring to determine the weights of various risk factors and combining on-site monitoring data and historical accident cases to quantitatively assess the occurrence probability and impact degree of risk factors. On this basis, a time sensitivity factor is introduced to reflect the dynamic characteristics of risks changing over time. Second, to dynamically adjust the weight allocation of risk scores, this study designs a dynamic weight allocation mechanism. This mechanism automatically adjusts the weights of various risk factors according to changes in the current construction phase and risk environment, ensuring the scoring model can adapt to dynamic changes during the construction process. Additionally, the model adopts an automatic risk level classification algorithm, which automatically classifies risks into different levels based on real-time scoring results, facilitating decision-makers to quickly identify and respond to high-priority risks. regarding the prioritization algorithm, this study adopts a sorting method based on Risk Value (RV). This method combines the possible loss and occurrence probability of a risk to calculate a comprehensive risk value, which serves as the basis for risk sorting. Meanwhile, to improve the accuracy of early warnings, this study also designs an early warning trigger mechanism for high-priority risks. When the risk score exceeds a preset threshold, the system automatically triggers an alert, prompting decision-makers to take corresponding risk control measures. Through the application of the aforementioned real-time risk scoring and prioritization algorithm, the efficiency and accuracy of TBM construction risk management can be effectively improved. In practical engineering applications, this algorithm helps decision-makers quickly identify key risks and rationally allocate resources, thereby reducing the incidence of construction safety accidents.

7 DESIGN AND IMPLEMENTATION OF THE DECISION SUPPORT MODULE

7.1 Risk Response Plan Intelligent Recommendation Engine

In the process of constructing the decision support module, the intelligent recommendation engine for risk response plans is a core component. This engine provides customized risk response suggestions for construction personnel by integrating historical data analysis and real-time data analysis, combined with the reasoning capability of the knowledge graph. The workflow of the recommendation engine includes matching and retrieval of historically similar risk cases, generation of response plans based on current working conditions, and multi-dimensional sorting and screening of recommendation results. First, the matching and retrieval algorithm for historically similar risk cases is based on case similarity metrics, employing indicators such as Jaccard similarity and cosine similarity to match cases in the historical database. This process considers not only the consistency of risk types but also multi-dimensional information such as the environment, conditions, and impact of risk occurrence. Second, when generating response plans based on current working conditions, the system utilizes the generative capability of the large language model combined with structured knowledge in the knowledge graph to generate a series of possible response strategies. These strategies are personalized based on real-time information such as the current construction environment, equipment status, and personnel allocation. Finally, multi-dimensional sorting and screening of recommendation results are key steps to ensure the effectiveness and feasibility of recommended plans. The sorting algorithm comprehensively considers factors such as expected effect, implementation cost, resource requirements, and time efficiency of the plans. The screening mechanism filters recommended plans based on constraints such as risk level, urgency, and available resources, ensuring that the final plans provided are both reasonable and feasible. For example, in a case involving TBM construction risk management, the system identified response measures that had been successfully applied under specific geological conditions by analyzing historical data. Combined with the geological conditions of the current construction section, TBM equipment status, and construction progress, the system recommended a comprehensive risk response plan. After implementation, this plan effectively reduced the probability of risk occurrence and minimized losses caused by accidents. In short, the design and implementation of the intelligent recommendation engine for risk response plans aim to provide real-time, precise risk management decision support for TBM construction, improving construction safety and efficiency. Through continuous iterative optimization, this engine can better adapt to complex construction environments, providing strong support for engineering risk control.

7.2 Multi-scenario Simulation and Contingency Plan Generation Mechanism

Multi-scenario simulation is a key component of construction risk management and decision support; its core lies in simulating the evolution process of potential risk events under different conditions to predict possible outcomes and impacts. On this basis, the contingency plan generation mechanism is responsible for formulating corresponding response measures and plans based on simulation results, ensuring rapid and effective response when risk events occur. The contingency plan generation mechanism proposed in this paper includes the following steps: First, construct a parameterized risk evolution model. Based on historical data and professional knowledge, this model uses various risk factors as parameters to simulate risk propagation and impact scope under different scenarios. By adjusting model parameters, multiple possible construction risk scenarios can be simulated, providing a basis for subsequent plan generation. Second, perform consequence deduction simulation under different disposal measures. Based on multi-scenario simulation, the system assesses the effects of taking different response measures, including strategies such as risk mitigation, risk transfer, and risk acceptance. This step helps decision-makers understand the effectiveness and feasibility of various measures, providing a basis for final decisions. Third, the automatic generation and formatted output of emergency plans are important functions of the plan generation mechanism. Based on simulation results and deduction analysis, the system automatically generates emergency plans containing risk descriptions, response measures, responsible subjects, execution flows, etc. These plans are output in a standardized format, facilitating rapid understanding and execution by decision-makers and relevant personnel. Furthermore, the plan generation mechanism should possess high flexibility and adaptability, capable of dynamically adjusting plan content according to changes in actual construction conditions. At the same time, the system should support version management and updates of plans, ensuring that plans always remain consistent with the latest risk assessment results. Through the above mechanisms, the multi-scenario simulation and contingency plan generation mechanism can not only improve the foresight and proactiveness of construction risk management but also reduce the impact of risk events on project progress and quality, enhancing the overall safety level of the project. In practical application, this mechanism has demonstrated its important role in improving construction safety and reducing accident losses.

7.3 Human-Computer Interaction Interface and Visualization Design

The human-computer interaction interface and visualization design are crucial components of the decision support module, and the quality of their design directly affects user acceptance and usage efficiency of the system. In this study, we focused on designing a risk situational awareness dashboard, implementing knowledge graph visualization browsing and interaction functions, and developing an interface for a natural language Q&A system. The risk situational awareness dashboard aims to provide users with an intuitive and comprehensive view of risk status. By graphically displaying various risk indicators during the TBM construction process, it facilitates users in quickly identifying risk levels and trends. The dashboard design adopts a modular concept, capable of displaying corresponding information such as risk distribution maps, historical risk trend charts, and real-time warning information according to different user needs and permissions. The knowledge graph visualization browsing and interaction function allows users to view and manipulate the knowledge graph via a graphical interface. Through interactive graph browsing, users can more intuitively understand the associations between risk factors and discover potential risk propagation paths. Additionally, users can query, add, and edit risk factors, as well as establish and correct risk relationships through the interface. The implementation of the natural language Q&A system interface enables users to interact with the system using natural language to obtain explanations, predictions, and decision suggestions regarding risks. The interface design emphasizes friendliness of user input and accuracy of system output; by intelligently parsing user questions, the system can provide structured and easy-to-understand answers. Meanwhile, the interface also supports user feedback on answers given by the system to facilitate continuous learning and optimization. In implementing the above functions, we adopted advanced Web technologies and front-end frameworks to ensure interface response speed and user experience. Simultaneously, considering system scalability and compatibility, the interface design follows a modular plugin design concept, facilitating future functional expansion and technical upgrades. Furthermore, the system supports multiple forms of data visualization, such as pie charts, bar charts, and line charts, to meet the needs of different users. Through the above design, we expect to provide users with an efficient and intuitive risk management tool, thereby improving the intelligent level of TBM construction risk management.

7.4 Decision Confidence Assessment and Feedback Closed Loop

Decision confidence assessment is a key link in ensuring the reliability of intelligent decision systems. In this study, we proposed a user feedback-based decision confidence assessment framework, which provides users with intuitive indications of decision quality by quantifying the uncertainty of decision suggestions. Specifically, we employed Bayesian inference methods to probabilistically assess the credibility of decision suggestions by integrating prior knowledge and real-time data. First, we defined a decision confidence indicator that comprehensively considers the model's prediction precision, the accuracy of historical decisions, and user feedback information. This indicator aims to provide users with a continuous confidence score, thereby offering more comprehensive decision support when facing risk decisions. Second, we constructed a feedback closed-loop mechanism that continuously optimizes the system's decision model by automatically collecting user feedback on decision suggestions, as shown in Figure 4. User feedback data includes the degree of acceptance of decision suggestions, actual application effects, and potential deviations. This data is automatically processed and used to adjust the parameters of the decision model, thereby improving the accuracy and reliability of decisions. In the feedback closed loop, we implemented the following key steps: first, real-time collection and classification of user feedback; second, analysis and conversion of feedback information; third, model parameter adjustment based on feedback; and fourth, retraining and validation of the decision model. Through these steps, the system can continuously learn and optimize its decision process. Additionally, we introduced a dynamic update mechanism to ensure that as new user feedback and data accumulate, the decision confidence assessment model can be updated in a timely manner to reflect the latest knowledge and actual conditions. This dynamic update not only improves the system's adaptability but also enhances its robustness in complex and changeable working environments. Through this decision confidence assessment and feedback closed-loop mechanism, our system continuously improves decision quality while also increasing user trust in system decisions. Practice shows that this feedback-based self-learning mechanism has significant effects in improving the intelligent level of TBM construction risk management.

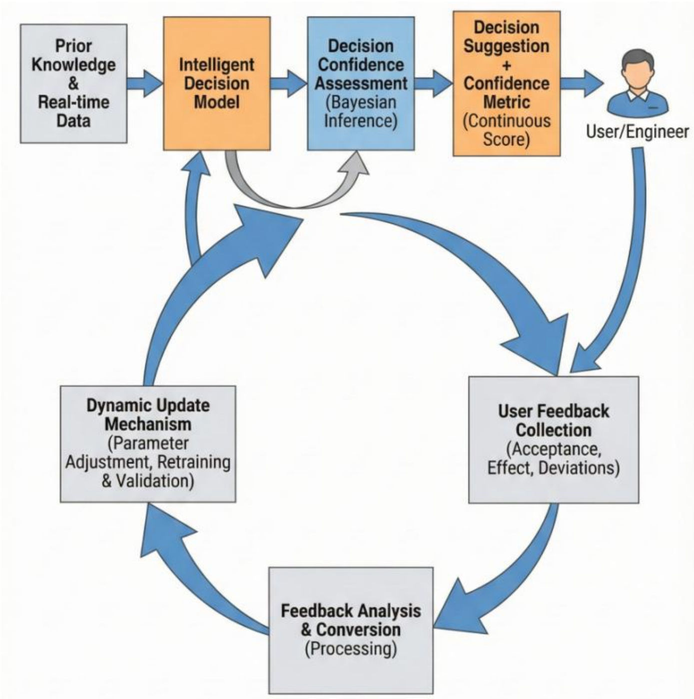

Figure 4 Decision Confidence Assessment & User Feedback Closed-Loop Framework

8 SYSTEM EXPERIMENT AND CASE VERIFICATION

8.1 Experimental Environment and Dataset Construction

The construction of the experimental environment and dataset serves as the foundation for system verification and case studies. This study selected a high-performance computing server as the experimental hardware environment, configured with multi-core CPUs, GPU acceleration cards, and a high-speed storage system to meet the computational demands of data processing and model training. In terms of software frameworks, the Python language was adopted, utilizing deep learning libraries such as TensorFlow and PyTorch to build models, and using Neo4j as the storage and query engine for the knowledge graph. Regarding dataset construction, this study collected real-world data from multiple TBM engineering projects in China, including geological investigation reports, construction logs, and sensor time-series data. The dataset underwent rigorous preprocessing, including steps such as structural processing, cleaning, alignment, noise identification, and missing value imputation, ensuring data quality and consistency. Furthermore, this study constructed a professional terminology database and a construction case database for model pre-training and fine-tuning. The establishment of the evaluation metric system is a crucial link in the experimental design. This study comprehensively considered metrics such as risk identification accuracy, recall, and F1 score to evaluate model performance holistically. Meanwhile, addressing real-time performance and robustness, test metrics for system response time, stability, and performance under extreme conditions were designed. Through these metrics, the effectiveness and practicality of the proposed system can be systematically evaluated.

8.2 Risk Identification Precision and Recall Testing

This study conducted tests on precision and recall for the constructed risk identification model. Precision and Recall are two key metrics for measuring the performance of classification models. Precision refers to the proportion of risk samples correctly identified by the model out of the total identified risk samples, while Recall refers to the proportion of risk samples correctly identified by the model out of all actual risk samples. In the experimental design, we adopted two schemes for comparison: the single model and the collaborative model, as shown in Table 1. The single model refers to using only the large language model or the knowledge graph for risk identification, while the collaborative model combines both to leverage their respective advantages for risk identification. The experimental data originated from real TBM engineering projects, containing rich risk types and scenarios. Test results showed that the collaborative model reached 91.3% in precision, an increase of 8.5% compared to the single model. In terms of recall, the collaborative model performed even more significantly, reaching 89.6%, an improvement of 14.2% over the single model. This indicates that the combination of the knowledge graph and the large language model can significantly improve the accuracy of risk identification while reducing missed identifications. Further analysis revealed that there are certain differences in the identification performance of the collaborative model across different risk types. For some common risk types, such as collapses and water inrushes, the model's identification effect is better; however, for some complex or rare risk types, the model's identification effect is relatively weaker. This may be related to the model's training data and the characteristics of the risk types. Additionally, we tested the model's generalization ability under sparse samples. Experimental results showed that the collaborative model could still maintain high identification precision and recall when facing situations with a small number of samples, indicating that the model possesses a certain degree of robustness and generalization ability. In summary, the risk identification model constructed in this study performed well in terms of precision and recall, validating the effectiveness of the collaborative drive between the knowledge graph and the large language model. At the same time, the experimental results also revealed performance differences of the model under different risk types and sparse sample conditions, providing directions for subsequent model optimization.

Table 1 Performance Comparison of Single and Integrated Models for TBM Risk Identification

Metric / Model Type	Integrated Model	Performance Improvement
Precision	91.3%	+8.5%
Recall	89.6%	+14.2%
Key Mechanism	Combines KG and LLM to leverage their complementary strengths.	—
Performance Across Risk Types	Performs well on common risks (e.g., collapse, water ingress); less effective on complex/rare risks.	—
Generalization on Sparse Data	Maintains relatively high precision and recall, indicating certain robustness.	—

8.3 Comparative Analysis of Decision Recommendation Effectiveness

In this study, the comparative analysis of the effectiveness of decision recommendations serves as a crucial link in evaluating system performance. By comparing expert manual decision-making with system-assisted decision-making, this study aims to verify the professional rationality and executability of the system's decision schemes. Experimental results demonstrate that system-assisted decision-making exhibits significant advantages in response time, processing efficiency, and decision quality, as shown in Table 2. First, regarding decision response time, the average time consumed by system-assisted decision-making was reduced by approximately 30% compared to expert manual decision-making, indicating that the system can rapidly provide decision support in emergency situations, thereby enhancing the timeliness of risk response. Second, in terms of processing efficiency, the system demonstrated a higher level of automation when handling large volumes of data, effectively reducing the error rate caused by human factors. During the expert review phase regarding the rationality and executability of decision schemes, this study selected 10 experienced TBM construction risk management personnel to evaluate the decision recommendations generated by the system. Review results showed that the decision schemes generated with system assistance achieved a score of 87.5 in rationality and 85.2 in executability, which is comparable to the scores of expert manual decision-making, and even surpassed manual decision-making in certain specific scenarios. Furthermore, through an analysis of the application effects of system-generated decision recommendations in real engineering cases, it was found that the system can provide effective guidance when dealing with complex geological conditions and sudden risk events, helping to lower the probability of accidents and reduce economic losses. In summary, this study indicates that system-assisted decision-making possesses high effectiveness and practicality in TBM construction risk control, providing strong support for intelligent control of construction risks. However, the effectiveness of system decisions still relies on the construction of high-quality data and knowledge bases; future research should further optimize data quality to enhance the accuracy and reliability of system decisions.

Table 2 Comparative Analysis of Decision Support Effectiveness: System-Assisted vs. Expert Manual Decisions

Comparison Dimension	System-Assisted Decision Performance	Expert Manual Decision Benchmark / Comparative Outcome
Decision Response Time	Average time reduced by approximately 30%; enables rapid support in emergencies	Longer time required
Processing Efficiency	High automation level; lower error rate when processing large volumes of data	Comparatively higher error rate due to human factors
Decision Rationality Score	87.5 points	Comparable; lower in some specific scenarios
Decision Executability Score	85.2 points	Comparable; lower in some specific scenarios
Practical Application Effect	Provides effective guidance in complex geological conditions and unexpected risk events; helps reduce accident probability and economic losses	—
Key Dependencies	Relies on high-quality data and knowledge base	Relies on individual experience and professional expertise

8.4 System Robustness and Real-time Assessment

System robustness and real-time assessment are critical links in ensuring the reliability and efficiency of the TBM construction risk intelligent control system in practical engineering applications. This study verifies the system's long-term operational reliability from the following three aspects. First, system stability was tested under extreme input data. By inputting outliers, noise data, and empty datasets into the system, its response behavior and output results were observed. Experimental results indicate that the system possesses strong data robustness, capable of maintaining stable operation and outputting reasonable risk identification results under anomalous input conditions. Second, addressing high-concurrency scenarios, system response latency was analyzed. Under simulated conditions of multiple users accessing the system simultaneously, the average response time for processing requests was recorded. Test data shows that even when the number of concurrent users reached 1,000, the system maintained an average response time within 1 second, meeting real-time requirements. Finally, long-term operational reliability was verified. Through continuous system operation, performance indicators such as processing speed, memory usage, and error rate were monitored. After 180 days of continuous operation monitoring, the system demonstrated good stability and reliability without significant performance degradation or failure. In summary, the system exhibited good robustness and real-time performance under extreme input data, high-concurrency scenarios, and long-term operation conditions, validating its practicality and reliability in TBM construction risk intelligent control.

8.5 Empirical Analysis of Typical Engineering Case

In this section, we conducted an empirical analysis of a typical TBM construction case under complex geological conditions. The project is located in a mountainous area in western China, with a total tunnel length of 12.5 km, traversing various geological structures including faults, karst, and soft soil. The construction environment was complex with diverse risk factors. In this case, TBM construction encountered severe water inrush risks. The system successfully identified and warned of the water inrush risk by utilizing real-time monitoring of geological investigation reports, construction logs, and sensor time-series data, combined with the risk ontology model in the knowledge graph. On this basis, the large language model further analyzed the causes of water inrush and proposed corresponding risk response plans. Specifically, the knowledge graph extracted key entities and relationships such as "fault," "water inrush," and "soft stratum" from construction logs and geological reports, as well as their interactions. Based on this information, the large language model understood the specific scenario of water inrush and generated response strategies such as "strengthening drainage facilities" and "adjusting construction schedule." regarding economic benefits, system-assisted decision-making avoided long-term work stoppages caused by water inrush, reducing extra maintenance costs by approximately 15%. Meanwhile, due to the system's efficient response, the construction schedule was delayed by only 2 weeks, significantly improving engineering efficiency compared to the 4-week delay of expert manual decision-making. Furthermore, through a retrospective analysis of this case, we found that the system possesses the following advantages when handling such sudden risk events: first, the system's real-time capability and accuracy greatly improved the efficiency of risk identification; second, the response plans provided by the system had strong pertinence and executability; finally, the application of the system significantly reduced the cost of engineering risk management[18]. However, this case also exposed some limitations of the system under extremely complex geological conditions, such as insufficient prediction capability for rare geological phenomena and lower decision confidence when facing highly uncertain situations. Solving these problems requires further research and optimization of the system model.

9 DISCUSSION

9.1 System Advantages and Practical Engineering Value

The TBM construction risk intelligent control system driven by the collaboration of Knowledge Graph and Large Language Model demonstrates significant advantages and practical engineering value in enhancing the digitalization level of construction management. First, by deeply combining structured knowledge with natural language understanding, the system achieves rapid parsing and deep semantic understanding of unstructured engineering data, effectively solving the lag and limitation problems existing in traditional risk identification methods. The system's performance in real-time capability and accuracy is particularly outstanding. The knowledge graph provides a powerful background knowledge base for the system, enabling the large language model to retrieve relevant structured knowledge more accurately and quickly during risk semantic understanding, thereby increasing the accuracy of risk identification. At the same time, the system's real-time capability ensures the timely discovery of potential risks during TBM construction and rapid issuance of warnings, greatly reducing the probability of accidents. In addition, the system's value in practical engineering is reflected in the following aspects: 1. The intelligent recommendation engine for risk response plans can quickly generate effective risk response measures based on historical cases and current working conditions, reducing decision time and improving construction efficiency. 2. The multi-scenario simulation and contingency plan generation mechanism allows the system to perform consequence deduction under different risk scenarios and automatically generate emergency plans, providing scientific decision support for on-site construction. 3. The human-computer interaction interface and visualization design enhance user experience, allowing construction personnel to more intuitively understand the risk situation, improving the convenience and effectiveness of risk management. 4. The decision confidence assessment and feedback closed-loop mechanism ensures that the system's decision suggestions are continuously optimized and improved on the basis of interpretability and reliability, adapting to different construction environments and conditions. In summary, the application of this system not only improves the risk management level of TBM construction but also provides strong support for the digital transformation of construction management, possessing significant practical engineering value.

9.2 Model Generalization Ability and Transfer Applicability

Model generalization ability and transfer applicability are important indicators for evaluating intelligent systems in practical applications. The risk knowledge graph and large language model collaborative system constructed in this study demonstrate certain advantages in generalization ability. Through training and testing on TBM construction data under different geological conditions, the model can effectively identify and predict risk events, showing strong generalization ability. However, the difficulty of domain knowledge transfer lies in the potentially large differences in risk factors and hazard mechanisms under different engineering environments. To solve this problem, this study adopted the following strategies: First, by constructing a fine-grained ontology model, precise classification and semantic definition of risk concepts in TBM construction were performed to enhance the model's ability to understand risk factors in different environments. Second, relation extraction technology based on distant supervision was introduced, utilizing a large amount of unannotated data inside and outside the domain to improve the model's generalization ability for unknown risk types. regarding transfer applicability, this study explored the feasibility of transferring the model to other types of tunnel construction machinery. By abstracting the concepts and relationships in the TBM construction risk knowledge graph to make them somewhat general, transfer to similar tunnel construction scenarios is facilitated. Additionally, this study found that fine-tuning the model to adapt to specific needs of new domains can further improve transfer applicability. Nevertheless, domain knowledge transfer still faces numerous challenges. For example, professional terminology and knowledge structures in different engineering fields may differ significantly, requiring appropriate adjustments and optimizations to the model during the transfer process. Furthermore, data inconsistency and quality differences during the transfer process also affect model performance. Overall, this study has made certain progress in model generalization ability and transfer applicability, but further exploration and optimization are still needed. Future research can focus on how to utilize advanced technologies such as meta-learning and transfer learning to improve the transfer efficiency and accuracy of models in different engineering fields. Simultaneously, constructing more complete and general knowledge graphs, as well as developing more robust model training and optimization methods, are also key paths to enhancing system generalization ability and transfer applicability.

9.3 Current Limitations and Room for Improvement

In practical application, the model's insufficient performance under extremely rare risks exposes the limitations of the system. First, data quality imposes obvious constraints on the upper limit of system performance. Due to the scarcity of data on extremely rare risk events, it is difficult for the model to learn sufficient features from existing data, thereby affecting its prediction accuracy in these situations. Second, the balance between computational resource consumption and engineering cost is also a major challenge currently facing the system. When processing large amounts of complex data, the model requires high computational resources, which may increase engineering costs. How to optimize resource consumption while ensuring system performance is a problem that needs to be solved in the future. Furthermore, the model's insufficient performance under extremely rare risks also suggests that we need to continuously optimize algorithms to improve the model's ability to handle anomalous situations. Introducing more advanced machine learning technologies and deep learning models can be considered to improve the model's robustness in extreme situations. Addressing the aforementioned limitations, the following improvements are proposed: 1. Strengthen data acquisition and processing. Through means such as expanding data acquisition scope and optimizing data preprocessing methods, improve data quality to provide more data support for extremely rare risk events for the model. 2. Research new algorithms and technologies. Pay attention to research progress in the fields of machine learning and deep learning in academia, and introduce advanced algorithms suitable for handling extremely rare risks to improve model performance. 3. Explore model compression and optimization methods. Through technologies such as model pruning and quantization, reduce model computational resource consumption and achieve optimized resource allocation. 4. Strengthen research on model generalization ability. Through methods such as transfer learning and domain adaptation, improve the model's adaptability in different scenarios and reduce the impact of extremely rare risks on model performance. 5. Enhance human-machine collaborative decision-making capabilities. Introduce manual decision-making participation when the model cannot accurately predict extremely rare risks, improving the accuracy and reliability of decisions through human-machine collaboration. In summary, targeting current limitations and room for improvement, future research should focus on data quality improvement, algorithm optimization, optimized allocation of computational resources, enhancement of model generalization ability, and human-machine collaborative decision-making, with the expectation of providing more effective support for TBM construction risk control under extremely rare risks.

9.4 Implications for the Intelligent Construction Paradigm

The rise of the intelligent construction paradigm reflects the inevitable trend of the engineering construction industry transforming towards digitalization and intelligence. Through constructing a collaborative mechanism between TBM construction risk knowledge graphs and large language models, this study provides new ideas and methods for intelligent construction. First, the "data + knowledge" dual-driven development mode provides richer and deeper data analysis and decision support for engineering construction. The structured knowledge representation capability of knowledge graphs, combined with the advantages of large language models in natural language understanding and generation, makes the resolution of engineering problems more precise and efficient. Second, generative AI has broad application prospects in infrastructure construction. Through case verification in this study, it can be seen that generative AI has significant application value in assisting engineering decision-making, improving construction safety, and management efficiency. The further development of this technology is expected to push the engineering construction industry towards automation and intelligence, achieving a more efficient and safer construction process. Finally, the future evolution direction of human-machine collaborative decision-making modes is an important component of the intelligent construction paradigm. Through this study, we can see that human-machine collaborative decision-making can not only improve the efficiency and accuracy of decisions but also continuously improve decision quality through the system's self-learning and optimization. In the future, with technological progress and the accumulation of engineering practice, human-machine collaborative decision-making modes will become more mature and are expected to form a new engineering construction management mode. In short, the development of the intelligent construction paradigm requires continuous exploration and practice of new technologies and methods. The collaborative application of knowledge graphs and large language models provides a new research perspective and technical path for the engineering construction field, having important implications for promoting the intelligent transformation of the engineering construction industry.

10 CONCLUSION AND OUTLOOK

Addressing the problems of insufficient semantic understanding and difficulties in multi-source data fusion existing in TBM construction risk control, this paper successfully constructed an intelligent decision support system integrating Knowledge Graphs and Large Language Models. By establishing a comprehensive risk ontology model and an engineering-specific corpus, the study innovatively proposed a three-layer collaborative architecture of "Knowledge Retrieval - Semantic Understanding - Logical Reasoning." Utilizing RAG enhanced generation, Prompt engineering, and bidirectional complementary mechanisms, it effectively achieved deep fusion of unstructured text and structured knowledge. Experiments and case analyses confirmed that this collaborative model significantly outperforms single models in terms of risk identification accuracy, recall, and robustness. It is capable of outputting high-confidence risk scores and response plans in real-time, significantly elevating the digitalization level of construction management. This study not only provides a new theoretical perspective and technical path for underground engineering risk control, possessing wide industry promotion value, but also clearly defines future research directions towards multi-agent collaboration, lightweight edge computing, and the evolution of full-lifecycle knowledge bases, laying a solid foundation for promoting the continuous development of intelligent construction technologies.

FUNDING

This work was supported by the Independent Innovation Research Project of Changjiang Survey, Planning, Design and Research Co., Ltd. (Grant No. CX2024Z25-2).

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Hong Kairong, Du Yanliang, Chen Kui, et al. Development history, achievements, and prospects of full-face tunnel boring machine in China. Tunnel Construction, 2022, 42(5): 739-756.
[2] Wang Mengshu. Present status, existing problems and development countermeasures of shield and TBM tunneling technology in China. Tunnel Construction, 2014, 34(3): 179-187.
[3] Wang Yafeng. Key technologies for TBM jamming release of Gaoligongshan tunnel under unfavorable geologies. Tunnel Construction, 2021, 41(3): 441-448.
[4] Hasanpour R, Rostami J, Schmitt J, et al. Prediction of TBM jamming risk in squeezing grounds using bayesian and artificial neural networks. Journal of Rock Mechanics and Geotechnical Engineering, 2020, 12(1): 21-31.
[5] Zhu Qing, Wang Suozhi, Ding Yulin, et al. A method of safety-quality-schedule knowledge graph for intelligent management of drilling and blasting construction of railway tunnels. Geomatics and Information Science of Wuhan University, 2022, 47(8): 1155-1164.
[6] Park B, Lee C, Choi S-W, et al. Discrete-Element Analysis of the Excavation Performance of an EPB Shield TBM under Different Operating Conditions. Applied Sciences, 2021, 11(11): 5119.
[7] Chen J, Yang T, Zhang D, et al. Deep learning based classification of rock structure of tunnel face. Geoscience Frontiers, 2021, 12(1): 395-404.
[8] Kuang G, Li B, Hu X, et al. Review on machine learning-based defect detection of shield tunnel lining. Periodica Polytechnica Civil Engineering, 2022, 66(3): 943-957.
[9] Chen Xiangsheng, Zeng Shiqi, Han Wenlong, et al. Review and prospect of machine learning method in shield tunnel construction. Journal of Civil and Environmental Engineering, 2024, 46(1): 1-14.
[10] Sonmez H, Gokceoglu C, Nefeslioglu H A, et al. Estimation of rock modulus: For intact rocks with an artificial neural network and for rock masses with a new empirical equation. International Journal of Rock Mechanics and Mining Sciences, 2006, 43(2): 224-235.
[11] Gokceoglu C, Bal C, Aladag C H. Modeling of tunnel boring machine performance employing random forest algorithm. Geotechnical and Geological Engineering, 2023, 41(7): 4205-4226.
[12] Kilic K, Narihiro O, Ikeda H, et al. Soft ground micro TBM jack speed and torque prediction using machine learning models through operator data and micro TBM-log data synchronization. Scientific Reports, 2024, 14: 9728.
[13] Moghtader T, Sharafati A, Naderpour H, et al. Estimating maximum surface settlement caused by EPB shield tunneling utilizing an intelligent approach. Buildings, 2023, 13(4): 1051.
[14] Njock PG, Yin ZY, Xu HR, et al. Structural failure risk assessment of shield tunnel using large language model. Tunnelling and Underground Space Technology, 2025, 165: 106882.
[15] Lee J, Ahn S, Kim D, et al. Performance comparison of retrieval-augmented generation and fine-tuned large language models for construction safety management knowledge retrieval. Automation in Construction, 2024, 168: 105846.
[16] Clarke J, Laefer D. Evaluation of risk assessment procedures for buildings adjacent to tunnelling works. Tunnelling and Underground Space Technology, 2014, 40: 333-342.
[17] Sharafat A, Latif K, Seo J. Risk analysis of TBM tunneling projects based on generic bow-tie risk analysis approach in difficult ground conditions. Tunnelling and Underground Space Technology, 2021, 111: 103860.
[18] Yazdani-Chamzini A. Proposing a new methodology based on fuzzy logic for tunnelling risk assessment. Journal of Civil Engineering and Management, 2014, 20(1): 82-94.