Application of Data Mining Technology in Civil Engineering Systems

Yu Xiu

{13945823636@163.com}

School of Environmental Arts&Architectural Engineering, Heilongjiang University Of Technology, Jixi, Heilongjiang, China

Abstract: The level of science and technology and management is gradually advancing, and the management of engineering projects is gradually shifting from traditional management mode to modern management mode. Data mining (DM) technology is increasingly widely used in engineering project construction, but currently, exploration in this area is still in the initial stage, lacking corresponding theoretical basis and practical experience to guide practice. This article first introduced the theoretical development status of DM, then analyzed the main problems that need to be solved in the decision-making process of civil engineering systems. Finally, with examples, it explained how to use neural network algorithms and support vector machine models in DM technology to analyze, organize, and extract engineering information, in order to optimize design schemes and improve engineering quality. A comparison was made between the traditional method and the DM technology based health status detection of bridge construction structures. The experimental data showed that the evaluation results based on DM technology were superior to traditional methods, and the accuracy of the detection results was improved by about 4.43%. This can effectively achieve the engineering cost management goals and provide a reliable basis for project investment decision-making. In the design and construction process of building structures, DM technology can monitor and analyze the design and construction process in real-time, and automatically adjust plans or modify measures based on data information, laying a solid foundation for achieving reasonable design and efficient construction. At the same time, it also provides useful contributions for people to understand the actual situation and predict the future.

Keywords: Data Mining, Civil Engineering, Neural Networks, Support Vector Machines, Engineering Quality

1. Introduction

Civil engineering, as a special construction technology, its construction quality is related to the competitive advantage and development speed of the entire construction enterprise and even the country in the market. At present, society has increasingly high requirements for construction projects, and quality issues that arise during the construction process of civil engineering are also common. Therefore, how to

ICSETPSD 2023, November 17-18, Coimbatore, India

DOI 10.4108/eai.17-11-2023.2342713

effectively control the quality of engineering has become an important issue that construction units have to face. DM technology has strong analytical functions, and the use of DM technology for engineering project cost control and schedule control would achieve good results. This article would vigorously explore this new field, hoping to provide reference for relevant personnel.

As a promising Structural health monitoring tool, distributed optical fiber sensors can monitor the changes of environmental parameters of structures by connecting optical fiber cables to structures or embedding them into structures in civil engineering infrastructure [1]. The American Association of Civil engineer has prepared a practice manual for canal automation of irrigation systems, focusing on the technical aspects of modernization of water conservancy and irrigation systems in civil engineering by using automated channel control systems [2]. Surveying and on-site design are one of the most important parts of civil engineering project design, including the planning and implementation of infrastructure engineering and its layout. The application of mobile surveying and mapping systems can check and ensure the accuracy of the project [3]. A constructivist theoretical exploration was conducted on the formation of professional identity for students who were identified as disabled during their studies in civil engineering and their first year of work, emphasizing how social culture affects students' navigation in their undergraduate civil engineering careers, thereby promoting their integration into engineering [4]. In order to enhance students' problem-solving abilities in civil engineering and building management education, structured role-playing can be used as a teaching method to significantly enhance their problem-solving skills through planning during the project and communication with specific roles [5]. The use of planar Cartesian coordinates to ensure compatibility with measurements using electronic distance measurement allows for the use of static global navigation satellite system positioning for georeferencing precise measurement networks, and can be used for different applications in civil engineering, helping to evaluate their accuracy and accuracy [6]. For civil engineering systems, a good structure is crucial for building safety and economic development, and traditional methods are difficult to achieve these goals.

In recent years, DM technology has been widely used to solve various Structural health monitoring problems due to its powerful computing power. In terms of civil engineering structural damage identification, evolutionary technologies such as Particle swarm optimization algorithm, genetic algorithm and ant colony optimization algorithm have been applied as part of the process of achieving accurate solutions [7]. Rock burst is the extreme release of Strain energy stored in the surrounding rock, which may lead to casualties, damage to underground structures and equipment, and ultimately endanger the economic feasibility of the project. Three new DM technologies, neural network, decision tree and gene expression programming, have guiding significance for rock burst prediction [8]. Real-time high-precision positioning applications in relevant economic sectors such as precision agriculture, transportation, civil engineering or surveying and mapping need positioning and navigation services based on global navigation satellite systems, and big data architecture and data analysis based on DM algorithms is the best way to achieve this requirement [9]. DM technology provides more alternative solutions for construction enterprises, while also bringing new opportunities for civil engineering design and implementation work.

There are many problems in the construction and management process of traditional civil engineering, such as uncertain construction period and difficult quality control, which have led to significant safety hazards during the use of the project [10]. As an advanced information technology, DM technology can provide convenient and effective information services for people, thereby achieving the goal of reducing engineering costs and improving engineering quality [11]. Based on the neural network algorithm in DM technology and combined with the actual situation in the field of civil engineering, this article analyzed the application scope and characteristics of DM technology, and elaborated on its specific implementation methods in detail. Compared with existing research, neural networks based on DM have strong adaptability, fast learning speed, and high accuracy, and can be widely used in intelligent modeling, prediction, and decision-making in various complex scenarios.

2. DM Technology and Its Applications

With the continuous development of social economy and the increasing level of enterprise informatization, various information systems are constantly emerging, which to some extent affects the application effect of traditional statistical methods. DM is precisely based on this situation and has been widely applied. It can discover problems and find patterns by analyzing a large amount of real and relevant data, thereby obtaining valuable information. Then, based on the relationships between various subsystems in the mining system and their logical connections with other parts, it can achieve management and optimization of the entire network system. This type of technical means can reveal the reasons for changes in things or phenomena, and has strong practicality and operability, facilitating dynamic monitoring. It has broad application prospects in fields such as engineering construction and transportation.

2.1 Opportunities and Challenges Faced by Civil Engineering Systems

The trend of economic globalization is becoming increasingly evident, and China has become the world's largest civil engineering market. The urbanization process is still accelerating, the demand for infrastructure construction is strong, and the number of engineering projects is in a state of explosion. These changes in the situation have provided unprecedented development opportunities for the civil engineering industry, but also brought enormous pressure [12-13].

Firstly, the quality control of construction projects is the prerequisite and foundation for ensuring the safety and service life of building structures. Secondly, engineering cost control is the key to ensuring the overall and social benefits of the project. Finally, improving construction efficiency and saving time require a large amount of information and data to support. Considering the above three aspects, only the establishment of a perfect civil engineering project management system can really achieve the requirements of high quality and high efficiency, and provide strong support and powerful help for the implementation of engineering projects [14].

2.2 Key Technologies of DM Application in Civil Engineering

2.2.1 Linear Regression Analysis

This is an effective method of modeling and predicting based on data, which identifies a certain connection between known and unknown values through statistical inference, and then proposes new possible results. It has strong adaptability and can provide accurate and reasonable explanations for complex problems, but its accuracy is not high.

2.2.2 Cluster Analysis

Cluster analysis is the application of similarity theory to study the laws of mutual correlation between things and make them a whole system, thus revealing their internal structure and evolution mechanism, and achieving a transformation from phenomenon to essence. This analysis method is more suitable for quantitatively describing the actual situation and can reduce the uncertainty caused by human judgment factors.

2.2.3 Neural Network

This type of technology is widely used in nonlinear information processing, mainly including BP (back propagation) neural network algorithm, genetic algorithm, support vector machine, pattern recognition and other methods. By using them to establish a model and solve the objective function, the required information can be obtained, and the computational complexity is small and the speed is fast, which can be adapted to multi-sensor fusion processing.

Civil engineering construction is a long-term and complex systematic project. To ensure the project quality, strict and scientific construction technology and management methods must be available, so as to effectively reduce the project cost, achieve the Sustainable Development Goals and maximize the economic benefits of enterprises [15-16].

3. Civil Engineering Structural Health Monitoring System

The use of computer technology can comprehensively assess the health status of engineering structures, which can be divided into two parts. One is the hardware equipment with on-site monitoring instruments as the core, and the other is the software system with databases as the core. The two form a unified information network system through sharing information with other relevant institutions and users through the network, providing people with a safe and efficient working environment. The specific architecture of the system is shown in Figure 1.

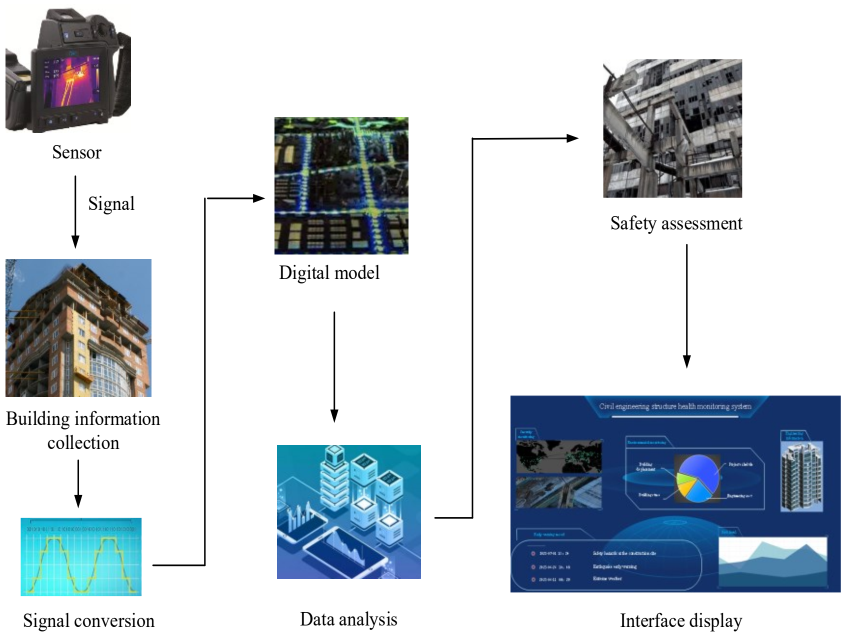

Fig.1 Architecture of Civil Engineering Structural Health Monitoring System

The civil engineering structural health monitoring system consists of four modules: sensors, data collectors, data analysis, and safety assessment. Among them, the sensing part mainly consists of various measuring devices and corresponding measurement and control equipment installed on various components inside the building, such as infrared thermal imagers, cameras, etc. [17]. The data collector is used to receive signals from sensors and convert them into digital signals for transmission to a microcontroller. After preprocessing, it is sent to a database to store relevant parameters, making it a complete digital model. Data analysis is mainly to reasonably organize, sort out and classify the original data obtained, generally including the comprehensive processing of information such as the displacement of buildings, or the stress that causes the deformation of building materials due to the action of external forces, so as to provide more accurate design basis for the engineering management department. The final safety assessment is the most critical and important part of the entire project process. Not only should structural damage identification be carried out, but also various environmental factors should be combined to analyze the safety issues of building structures. The entire system has strong adaptability and anti-interference ability to changes in environmental conditions, which can meet on-site construction requirements and improve detection efficiency [18-19].

4. Civil Engineering Structural Health Detection Based on DM Technology

Assuming there is a training set A for structural health monitoring in construction projects, and $A = \{a_1, a_2, \dots a_n\}$. F is the mapping function for kernel principal component analysis, satisfying the following conditions:

$$\sum_{i-1}^{I} F(a_i) = 0 \quad (1)$$

Where, the value range of i is $[1, I]$, and the covariance matrix is established:

$$B = \frac{1}{I}\sum_{j-1}^{I} F(a_j)F(a_j)^T \quad (2)$$

By Eigendecomposition of a matrix of Formula (2), it can be obtained:

$$\mu q = Bq \quad (3)$$

Among them, μ is the characteristic value; q is the feature vector, and $\mu \geq 0$.

$$q = \{F(a_1), F(a_2), \dots, F(a_n)\} \quad (4)$$

So, the equivalent form can be obtained:

$$\mu(F(a_i), q^r = (F(a_i), Bq^r) \quad (5)$$

$$q^r = \sum_{i-1}^{I} B_i^r F(a_i) \quad (6)$$

A linear combination of 1 and 2. Then, the support vector machine is used to find the Hyperplane with the largest boundary, that is:

$$M_x(eN_x + v) - 1 \geq 0 \quad (7)$$

By converting to the Lagrand problem, it can be obtained:

$$L(e, v, s) = \frac{1}{2}\|e\|^2 + \sum_{x=1}^{y} s_x[M_x(eN_x + v) - 1] \quad (8)$$

Among them, s_x is the Lagrand coefficient, and $x = 1,2, \dots, y$.

The classification decision function of the final support vector machine can be expressed as:

$$G = \mathrm{sgn}\left[\sum_{x=1}^{y} M_x s_x(N \cdot N_x) + v\right] \quad (9)$$

Among them, M_x is the output, and N_x is the input; e and v are weight and bias vectors, respectively.

Based on DM technology, the health status of engineering structures is detected, and the extracted time-domain statistical features of engineering structures are normalized to obtain:

$$G_x' = \frac{G_x - min(G_x)}{max(G_x) - min(G_x)} \quad (10)$$

Among them, G_x is the original feature. The health status characteristics of engineering structures are taken as the input of the support vector machine, and the health status of engineering structures is taken as the output. Through the fitting of the relationship between the two, the parameters of the support vector machine are determined using Particle swarm optimization algorithm, and the health detection model of civil engineering structures can be established [20].

5. Evaluation of the Application Effect of DM Technology in Civil Engineering Systems

5.1. Experimental Design

In this paper, the bridge buildings in a certain area were taken as the experimental objects. A total of 100 samples were selected, and the damage degree of their building structures was divided into four grades, namely, mild damage, moderate damage, serious damage and chemical damage. Among them, chemical damage was the most serious, and showed obvious corrosion phenomenon, such as the erosion of PVC combustion gas on concrete structures. The classification criteria and sample distribution for the four levels are shown in Table 1.

Table 1. Classification Criteria and Sample Distribution of Damage Severity Levels

Level	Damage degree	Damage index	Sample size
I	Mild injury	<0.1	50
II	Moderate injury	0.1~0.4	30
III	Serious injury	0.4~0.7	15
IV	Chemical damage	0.7~1	5

5.2. Data

In order to evaluate the application effect of DM technology in civil engineering systems, traditional methods, kernel principal component analysis based methods, and support vector machine methods were used to detect the health status of sample bridge building structures. The detection quality and efficiency were evaluated from two aspects: accuracy and detection time, and the data was recorded and analyzed.

5.2.1 Accuracy

The accuracy rate can determine whether the two technical methods can achieve correct identification and determine the type of damage, as shown in Figure 2.

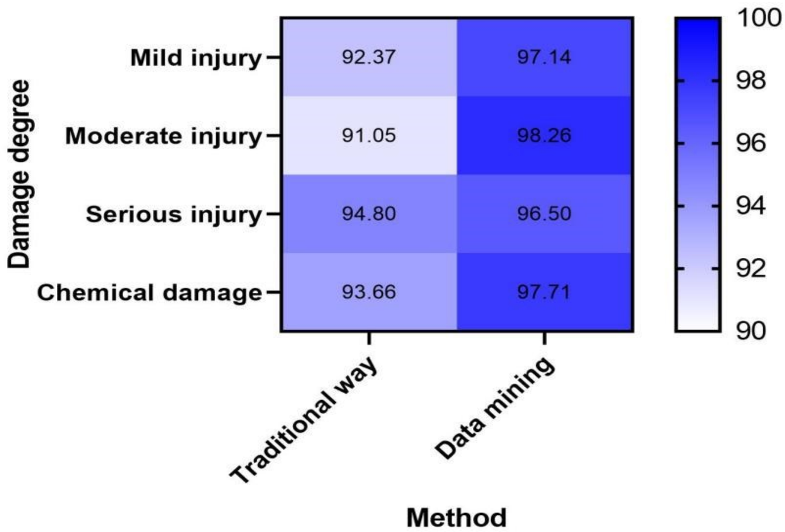

Fig.2 Accuracy of two technologies in detecting the health status of bridge engineering structures

The horizontal axis in Figure 2 represents traditional and DM techniques, respectively. The vertical axis represents the detection accuracy of each type of damage degree. The heavier the color, the closer the numerical value is to 100, and the higher the accuracy. From the data in Figure 2, it can be seen that the overall color of the data in the left column was relatively light, and the accuracy of the detection results was distributed between 90% and 95%; the data in the right column was darker in color, and the accuracy rate was basically 95% higher, with the highest even reaching around 98.26%. After calculation, it can be concluded that the average accuracy of traditional methods was about 92.97%, while the average accuracy based on DM technology was about 97.4%, which was an improvement of about 4.43% compared to traditional methods. Therefore, DM technology can more accurately and quickly obtain more information and perform effective analysis and processing compared to traditional methods, thereby improving the recognition accuracy of abnormal structural damage features.

5.2.2 Detection Time

The time and average value spent on completing the detection using two methods were statistically analyzed and calculated, as shown in Figure 3.

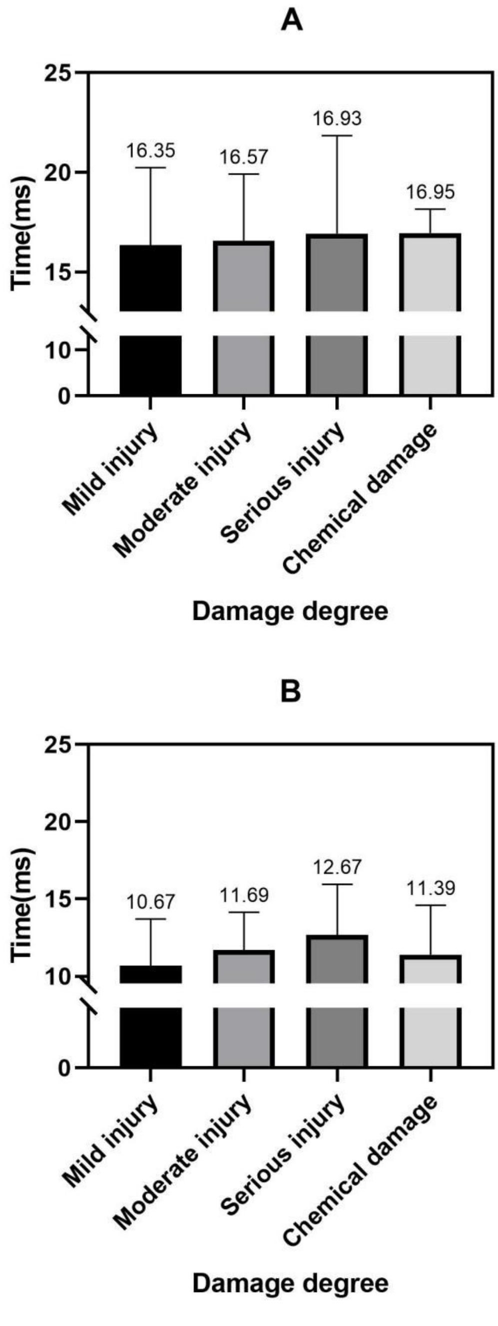

Fig.3 Detection time of two technologies in the health status detection of bridge engineering structures

A. The detection time of traditional methods in the health status detection of bridge engineering structures

B. The detection time of DM technology in the health status detection of bridge engineering structures

The horizontal axis in figures 3A and 3B represent four different degrees of damage, while the vertical axis represents the detection time. Taking the average detection time for each degree of damage, the average values obtained by traditional detection methods were around 16.35 milliseconds, 16.57 milliseconds, 16.93 milliseconds, and 16.95 milliseconds, respectively. The average detection time based on DM technology was around 10.67 milliseconds, 11.69 milliseconds, 12.67 milliseconds, and 11.39 milliseconds. From these data, it can be seen that traditional detection methods have significant limitations, while DM technology can improve the sensitivity of engineering detection to a certain extent, making it more efficient.

6. Conclusion

Civil engineering, as a special construction technology, occupies an important position in modern architecture. By strengthening the foundation structure, the overall strength and stability of the building can be improved. Applying DM technology to engineering practice can help optimize engineering design and construction quality, while also reducing engineering costs and promoting sustainable development of the construction industry. This article analyzed and discusses the problems existing in the current civil engineering construction process, and combined DM method and related algorithm theory as guidance to explain the importance of mining data information in practical work and specific implementation plans. It is hoped that this can further promote the improvement of project construction level and achieve healthy and rapid development of the social economy.

Acknowledgement

This work was supported by the Natural Science Foundation of China (No.61709882), the Natural Science Foundation Project of Jiangsu Province (No.2018A2039)

References

[1] Kenichi Soga, Linqing Luo. "Distributed fiber optics sensors for civil engineering infrastructure sensing." Journal of Structural Integrity and Maintenance 3.1 (2018): 1-21.

[2] Brian Wahlin, Darell Zimbelman. "Canal automation for irrigation systems: American society of civil engineers manual of practice number 131." Irrigation and Drainage 67.1 (2018): 22-28.

[3] Omar Al-Bayari. "Mobile mapping systems in civil engineering projects (case studies)." Applied Geomatics 11.1 (2019): 1-13.

[4] Cassandra McCall, Ashley Shew, Denise R. Simmons, Marie C. Paretti, Lisa D. McNair. "Exploring student disability and professional identity: Navigating

sociocultural expectations in US undergraduate civil engineering programs." Australasian Journal of Engineering Education 25.1 (2020): 79-89.

[5] Jingxiao Zhang, Haiyan Xie, Hui Li. "Improvement of students problem-solving skills through project execution planning in civil engineering and construction management education." Engineering, Construction and Architectural Management 26.7 (2019): 1437-1454.

[6] Nixon A. Correa-M., Liliana A. C.-C. "Precision and accuracy of the static GNSS method for surveying networks used in Civil Engineering." Ingenieria e Investigacion 38.1 (2018): 52-59.

[7] Meisam Gordan, Zubaidah Binti Ismail, Hashim Abdul Razak, Khaled Ghaedi, Haider Hamad Ghayeb. "Optimization-based evolutionary data mining techniques for structural health monitoring." Journal of Civil Engineering and Construction 9.1 (2020): 14-23.

[8] Roohollah Shirani Faradonbeh, Abbas Taheri. "Long-term prediction of rockburst hazard in deep underground openings using three robust data mining techniques." Engineering with Computers 35.2 (2019): 659-675.

[9] Angel M., Raquel Maria Capilla, Ana B. Anquela. "Big data architecture and data mining analysis for market segment applications of differential global navigation satellite system (GNSS) services: case study of the analysis of the demand for navigation and agriculture." The Journal of Navigation 75.2 (2022): 421-436.

[10] Jan Knippers, Cordula Kropp, Achim Menges, Oliver Sawodny, Daniel Weiskopf. "Integrative computational design and construction: Rethinking architecture digitally." Civil engineering design 3.4 (2021): 123-135.

[11] Priyank Sunhare, Rameez R. Chowdhary, Manju K. Chattopadhyay. "Internet of things and data mining: An application oriented survey." Journal of King Saud University-Computer and Information Sciences 34.6 (2022): 3569-3590.

[12] Danish Kazmi, David J. Williams, Mehdi Serati. "Waste glass in civil engineering applications—A review." International Journal of Applied Ceramic Technology 17.2 (2020): 529-554.

[13] Junping Zheng, Maoshan Qiang. "Understanding the changes in construction project managers' competences through resume data mining." Journal of Civil Engineering and Management 28.4 (2022): 305-319.

[14] Pengcheng Jiao, Amir H. Alavi. "Artificial intelligence in seismology: advent, performance and future trends." Geoscience Frontiers 11.3 (2020): 739-744.

[15] M. T., Peter M. "Utilizing drone technology in the civil engineering." Selected Scientific Papers-Journal of Civil Engineering 14.1 (2019): 27-37.

[16] Chen Wang, Yutong Tang, Mukhtar A. Kassem, Heng Li, Bingqing Hua. "Application of VR technology in civil engineering education." Computer Applications in Engineering Education 30.2 (2022): 335-348.

[17] Liqiang Xie, Tianxian Wang, Jianchun Xing, Xifeng Zhu. "An embedded surface acoustic wave pressure sensor for monitoring civil engineering structures." IEEE Sensors Journal 18.13 (2018): 5232-5237.

[18] Anshul Sharma, Pardeep Kumar, Hemant Vinayak, Raj KUMAR Patel. "A Review of Modeling and Data Mining Techniques Applied for Analyzing Steel Bridges." International Journal of Software Computing and Testing 7.1 (2021): 1-15.

[19] M. Gordan, Saeed R. Sabbagh-Yazdi, Z. Ismail, Kh. Ghaedi, H. Hamad Ghayeb. "Data mining-based structural damage identification of composite bridge using support vector machine." Journal of AI and Data Mining 9.4 (2021): 415-423.

[20] Nicolas ManziniORCID, A. Orcesi, Christian Thom, Marc-Antoine Brossault, Serge Botton, Miguel Ortiz, et al. "Performance analysis of low-cost GNSS stations for structural health monitoring of civil engineering structures." Structure and Infrastructure Engineering 18.5 (2022): 595-611.