A Framework on Quality Risk Early Warning for Hydraulic Engineering Construction Based on LSTM

Chunyong She¹, Ke Zhang², Jingchun Feng² and Le Cui²

¹ Zhejiang Water Conservancy and Hydropower Engineering Quality and Safety Management Center; Hangzhou, Zhejiang 310012
² Business School of Hohai University, Nanjing 211100, China

Abstract. With the reform of “strengthening management and service”, the focus of water conservancy project administration has gradually changed from the former mode of “attaching importance to examination and approval and ignoring supervision” to the mode of “attaching importance to service and strengthening supervision”. However, government supervision relies only on human resources, so it is inevitable that there will be overlapping supervision, loopholes, low efficiency and high cost. In this paper, the LSTM model in deep learning is used to predict the risks of water conservancy projects according to the text information such as quality supervision and inspection reports, and the risk early warning system of water conservancy projects is designed to provide auxiliary decision support for the quality supervision of water conservancy projects, which is of great significance to the formulation of quality supervision measures, allocation of supervision and management resources and the quality guidance of water conservancy projects.

1. Introduction

With more and more attention paid to the quality of water conservancy projects, strengthen the quality risk management has become one of the priorities of government departments. At present, in China, the quality supervision and inspection activities of water conservancy projects are mainly carried out by means of random inspection to check the hidden dangers of project quality and submit quality supervision and inspection reports according to regulations. The working mode tends to process monitoring and post punishment, and lacks comprehensive judgment and prediction of the risk status of water conservancy project quality. However, the construction process of water conservancy projects is irreversible and the risk factors during the construction are complex, so it is extremely important to improve the quality risk prevention ability. From the existing research basis, the quality risk management of water conservancy projects mainly analyzes the risk factors and evaluates the comprehensive risk level according to the actual characteristics of projects and different construction stages, further establishes the risk early warning system, and takes corresponding measures to reduce the risk losses and ensure the smooth construction of projects. Among them, risk assessment and early warning are mostly based on expert experience, and analytic hierarchy process, Monte Carlo method and other assessment methods are used to obtain the quality risk level of water conservancy projects. This process often needs to expend a lot of energy to organize expert evaluation activities, which is difficult to meet the requirements of timeliness and comprehensiveness of early warning work of water conservancy project quality risk.

Early warning of water conservancy project quality risk refers to the use of risk assessment methods to determine the future state of water conservancy project quality risk, and put forward preventive...
measures. The existing risk assessment and early warning methods of water conservancy projects can be divided into two categories: qualitative assessment method and quantitative assessment method. The qualitative assessment method realizes the risk assessment of water conservancy projects by non-quantitative means, mainly including Expert Scoring Method\cite{1}, Analytic Hierarchy Process\cite{2}, Risk Matrix Analysis\cite{3}, Fault Tree Analysis\cite{4}\cite{5}, etc. Quantitative assessment method refers to the establishment of mathematical models for the quantitative estimation of the risk level of water conservancy projects, mainly including Monte Carlo Method\cite{6}\cite{7}, Fuzzy Analysis\cite{8}, Grey System Theory\cite{9} and so on. In general, there are still some deficiencies in the research on quality risk assessment and early warning of water conservancy projects. On the one hand, the existing risk assessment and early warning methods of water conservancy projects generally obtain the basic data of risk factors through subjective quantification, ignoring the collection, processing and application of objective data sources (unstructured data, etc.). On the other hand, the existing risk assessment and early warning methods for water conservancy projects are mostly based on subjective experience evaluation. In practice, it is necessary to rely on the expert evaluation results to determine the weight of each risk factor, which tends to lead to uncertainty in the final risk measurement results.

In order to improve the accuracy and objectivity of risk early warning, it is necessary to build a data-driven risk assessment and early warning method for water conservancy projects. There are many theories that can be cited in data-driven risk assessment and early warning methods. Among them, the neural network model can deal with discrete and nonlinear large-scale risk assessment well, and it does not require subjective estimation of the weight of each risk factor, and it also has a certain research foundation. However, most of the input data of the existing neural network evaluation model are estimated manually, which is subjective to a certain extent and basically adopts a three-layer network structure. Studies have shown that the deepening of the number of layers in the neural network enables the model to better learn the characteristics of the input data, and the computational problems caused by the deepening of the number of layers in the network have been well solved in the field of deep learning. In view of the shortcomings of existing studies, this paper takes the quality supervision and inspection report as the objective input data, extracts the text characteristics of quality supervision as the input of the deep learning model, builds the technical framework of water conservancy project quality risk early warning, and improves the timeliness and accuracy of water conservancy project quality risk early warning.

2. Basic principles of deep learning
Deep learning generally refers to the use of artificial neural network model with multiple hidden layers to achieve complex classification learning tasks, so the study of artificial neural network is the basis of deep learning. Artificial neural network (ANN) mainly refers to the mathematical model inspired by biology and neurology. Figure 1 shows a typical neuronal structure. The input of neurons in the figure is the weighted sum of $x_1, x_2, \cdots, x_n$, and its calculation process is shown in Equation (1).

$$z = \sum_{i=1}^{n} w_i x_i + b$$

$$= W^T X + b$$

Where $W$ is the weight vector and $b$ is the bias. The weighted sum of input information is processed by nonlinear activation function to obtain neuron output as $y = f(z)$. The simplest nonlinear activation function includes step function and sigmoid function and so on.

There are three main types of common artificial Neural Networks: Feedforward Neural Networks (FNN), Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN). FNN model is the most common neural network structure, as shown in figure 1.
Feedforward neural network can be regarded as a function to realize complex mapping from input space to output space through simple nonlinear function. In feedforward neural network, each neuron belongs to different network layer. The neurons in each network layer can acquire the signals from the neurons in the upper layer and output them to the neurons in the latter layer after processing the signals. The initial network layer is called the input layer, the final network layer is called the output layer, and the other intermediate network layers are called the hidden layer. There is no feedback in the whole feedforward neural network, and the initial signal is one-way from the input layer to the output layer, which can be represented by a directed acyclic graph, as shown in figure 2.

In the feedforward neural network, the input of the neurons in the latter layer is the output of the neurons in the former layer, which can be expressed as

\[ z^l = W^T_l f_{l-1}(z^{l-1}) + b_l \]

Where \( z^l \) is the input of the \( l \) layer neurons, \( W_l \) is the weight vector of the \( l \) layer neurons, \( b_l \) is the bias of the \( l \) layer neurons, \( f_{l-1}(z^{l-1}) \) is the output of the \( l-1 \) layer neurons, and \( f \) represents the nonlinear activation function. Thus, the feedforward neural network can transmit information layer by layer to get the final output of the model \( Y \) (that is \( z^L \), \( L \) is the number of layers of the neural network). The whole network can be viewed as a composite function \( \phi(x, W, b) \), with \( x \) as the input of Layer 1 and \( z^L \) of layer L as the output of the whole function.

3. Technical framework for early warning of hydraulic engineering quality risks
The early-warning technical framework of hydraulic engineering quality risk from the perspective of deep learning can be divided into three parts: text processing, model construction and model application, as shown in figure 3.
3.1. Text processing

Text processing in this paper mainly focus on feature extraction from supervision text of water conservancy project quality. The text processing process of quality supervision mainly includes four steps: text collection, text cleaning, text segmentation and text vectorization. First, the original quality supervision text is cleaned, and abnormal data such as invalid text and duplicate text are processed to ensure the quality of input text. Then the text of quality supervision after cleaning is segmented. The word segmentation of quality supervision text is the basic work of quality supervision word vector training. Accurate word segmentation of quality supervision text provides guarantee for the feature extraction of subsequent quality supervision text. Text segmentation refers to the process of combining word sequences in a sentence into a sequence of words according to certain rules. Unlike English sentences, Chinese sentences lack natural boundaries between words and appear as word sequences. Therefore, Chinese word segmentation is more complicated and difficult than English word segmentation. Finally, in order to enable the computer to identify the quality supervision text features, word vector training is needed, and the acquired word vector is used to represent the quality supervision text features.
3.2. Model construction
The middle layer of the framework is the part of model construction, which is the core of the whole framework of water conservancy project quality risk early warning. Deep learning technology is used to construct the early warning model of water conservancy project quality risk by taking the text features of quality supervision as input and risk label data as output. Firstly, the model network layer is defined, that is, the number and connection mode of neurons in input layer, hidden layer and output layer are set. Then, taking the loss function as the model optimization objective, set parameters such as the initial network layer weight, learning rate, and training steps, train the weight parameters of each network layer, and obtain the optimal model according to the evaluation function.

3.3. Model application
The top layer of the framework is the model application section. The application of the model is mainly to evaluate the quality risk level of the classified quality supervision text according to the optimal model obtained.

4. Design of early warning system of water conservancy project quality risk

4.1. System design idea
The design of the early warning system of water conservancy project quality risk is to use the trained LSTM model to automatically and rapidly analyze the risk level of water conservancy project based on the project quality supervision report and the project data information, and to provide the decision support for the water conservancy project quality supervision department. According to the above requirements, the system adopts the Software architecture pattern based on B/S in the architecture design. The overall architecture design is shown in figure 4, which mainly includes the basic data layer, the core function layer and the actual business layer.
Figure 4. Overall architecture of early warning system for water conservancy project quality risk

4.2. Basic data layer
The basic data layer is the data foundation of system development. Its main function is to collect, integrate and process the data related to water conservancy project quality, and convert them into the data format required for risk rating, so as to facilitate the quality risk assessment model to invoke the relevant basic data. The data objects needed for water conservancy project quality risk assessment mainly include basic information of water conservancy project and text information of regular quality supervision and inspection report. The basic information of water conservancy projects includes construction years and historical maintenance records, etc., while the text of quality supervision includes inspection reports such as project name, existing problems and rectification opinions. The supervision and inspection report information is updated regularly. According to the basic project information and dynamically updated inspection information, the model of the LSTM deep learning can assess the risks of water conservancy projects more accurately and in real time.

4.3. Core functional layer
Risk assessment model is the core component of the whole system engineering. Its function is to program the complex risk grading standards in water conservancy projects and package them into "black box". After the text information is converted into standardized input, the project risk level can be automatically trained according to the input data. The LSTM model selected in this paper selectively controls the flow of information through the forget-gate structure, which determines how much of the previous historical information is retained until the current moment, and then combines the current memory and long-term memory to form a new cell state. The data redundancy in the time dimension can be reduced through the forgotten information. Moreover, LSTM structure is good at processing data with time sequence, so it is widely used in Text Analysis, Natural Language Processing, Time Series Prediction and other fields. Therefore, the LSTM model can be used to process the time series data of water conservancy project quality supervision and achieve better results. According to the calculation results of the model, early warning is given to high-risk water conservancy projects, and relevant personnel are reminded to take treatment measures to improve the project quality and ensure the safety of person and property.

4.4. Actual business layer
The early-warning system of water conservancy project quality risk is planned to be put into use in Zhejiang Water Resources Bureau, connecting the text report information of quality supervision and inspection in real time, and providing the interface of credit management system, effectively integrating information from multiple sources, and realizing risk rating of water conservancy project through the combination of automatic and human-machine interaction, so as to facilitate the quality supervision and inspection personnel to understand the status of project quality and safety. Change the current situation that the existing water conservancy project quality risk assessment method requires a lot of practice and human cost, and improve the timeliness and comprehensiveness of the water conservancy project quality risk assessment.

5. Summary and prospect
Based on the deep learning model of water conservancy project quality supervision and early warning demand, this paper designs and develops the deep learning-based early warning business system of water conservancy project quality of Zhejiang :

(1) By constructing a reasonable and effective water conservancy project quality risk assessment model, the future risks of the project can be predicted, and special monitoring and early warning can be carried out for the corresponding projects, which lays the foundation for the follow-up quality risk control and prevention strategies.
(2) Risk factors that affect the quality of water conservancy project are generated in each stage of project construction and are characterized by diversity and complexity. There are also differences in the degree of influence on the construction quality due to the interaction between the influencing factors of water conservancy project quality risk. It is difficult to directly estimate the quality risk of water conservancy projects through a single factor or a small number of factors. Therefore, it is necessary to construct the quality risk assessment model of water conservancy projects according to certain rules and carry out reasonable assessment on the whole project.

(3) According to the different risk levels predicted by the early warning system, it is of great significance for the quality risk management of water conservancy projects to allocate the required quality supervision and management resources, formulate corresponding quality supervision measures, and determine the control methods of different quality risk levels.

Through the water conservancy project quality risk early warning system designed in this paper, the visual display and management of the project quality supervision early warning information are realized, and the project risk is graded and evaluated by constructing the in-depth learning early warning model, realizing the all-round and whole-process accurate control of risk management, providing decision support and strong guarantee for the water conservancy project quality supervision. In the future, it will continue to improve the performance, scalability, stability and display efficiency of system kernel, optimize the risk early warning model, and further improve the business application capability.

References
[1] Xu Z H, Li S C, Li L P, Chen J and Shi S S 2011 Construction permit mechanism for karst tunnel based on dynamic risk assessment and control Journal of geotechnical engineering 33 1714-25
[2] Xu Z H, Li S C, Li L P, Hou J G, Sui B and Shi S S 2011 Risk assessment of water inrush and mud in karst tunnel based on analytic hierarchy process Rock and soil mechanics 32 1757-66
[3] Chen Z and Wang S J 2012 Research on the risks of water conservancy and hydropower projects based on hall three-dimensional structure Journal of shanxi university of finance and economics 34 30
[4] Ardeshir A, Amiri M, Ghasemi Y and Errington M 2014 Risk Assessment of Construction projects for water conveyor tunnels using Fuzzy fault tree analysis International Journal of Civil Engineering 12 396-412
[5] Zhou H B and Cai L B 2015 Risk and control of confined water in deep foundation pit engineering in soft soil area Journal of tongji university (Natural Science edition) 43 27-32
[6] Fan X E, Hu Z G and Jin P 2007 Integrated Risk Analysis of diversion System in Construction based on Monte-Carlo Method Progress in Water Science 18 604-8
[7] Marengo H, Arreguin F I and Romero I 2010 Risk assessment in hydraulic engineering projects: uncertainties and reliability J Tecnologia Y Ciencias Del Agua 1 5-35
[8] Zhou H B 2009 Fuzzy comprehensive evaluation method for deep foundation pit risk based on bayesian network Journal of Shanghai jiao tong university 43 1473-9
[9] Wang Y, Ma L, Bai J and Zhang Y X 2015 Evaluation of large International project Financial Risk Control based on Neural Networks Journal of Tongji University (Natural Science edition) 43 1104-10
[10] Pramanik S and Hussain A 2019 Text Normalization using Memory Augmented Neural Networks Speech Communication 109 15-23