

Research Article

Application and Comparison of Machine Learning Algorithms for Predicting Rock Deformation in Hydraulic Tunnels

Yixin Liu , Xuhua Ren , Jixun Zhang , and Yuxian Zhang

College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing, China

Correspondence should be addressed to Xuhua Ren; renxh@hhu.edu.cn

Received 6 March 2022; Revised 1 May 2022; Accepted 20 May 2022; Published 8 June 2022

Academic Editor: Xiang Peng

Copyright © 2022 Yixin Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Prediction of tunnel surrounding rock deformation is important for tunnel construction safety evaluation. In this paper, machine learning algorithms are used to carry out a comparative study of the surrounding rock deformation prediction. The applications of Gaussian process regression (GPR), support vector machine (SVM), and long short-term memory network (LSTM) in the prediction of surrounding rock deformation sequences are compared and analyzed. The actual data of a diversion tunnel in a southwest region are used as an example to evaluate and compare the single-step prediction model and multistep prediction model established by the above algorithm. The results show that the machine learning algorithm has good operation effect on the prediction of surrounding rock deformation. Overall, the SVM model has the best prediction effect and outperforms the other two algorithms in terms of tracking the trend of data changes and the degree of data fit.

1. Introduction

Rock deformation refers to the change of shape and volume of the rock around the underground cavern and the dislocation of the cavern wall. This phenomenon is generally caused when excavating underground cavities. And the amount of deformation of unreasonable surrounding rocks may trigger ground settlement and collapse, water gushing and rock explosion, and other engineering problems, eventually leading to the wreckage of underground tunnels, causing social losses, and casualties. Therefore, it is of great practical engineering significance to carry out real-time monitoring and safety evaluation of surrounding rock deformation during the construction process and operation of underground cavern projects. The traditional manual approach of building empirical models to solve problems based on measurement data suffers from significant limitations in experience and complexity, which can be remedied by machine learning's ability to capture data and learn from its trends. With the dramatic increase in computer performance, the efficiency of machine learning has also increased dramatically and is therefore widely used in various fields.

Many attempts have been made in the field of machine learning for surrounding rock deformation prediction. In terms of support vector machines (SVM), Goh et al. [1] demonstrated the predictive power of support vector machines in the mechanical properties of rocks. Yao et al. [2–4] did a lot of research on the application of support vector machine to predict the surrounding rock deformation, proposed SVM-based single-step prediction model and multistep prediction model for surrounding rock deformation, and gave an evaluation of the performance of these models and carried out optimization work. Shi et al. [5] established a prediction model for tunnel displacement settlement based on SVM and demonstrated the excellent performance of SVM in predicting tunnel surrounding rock deformation in conjunction with actual engineering. In terms of Gaussian process regression, Xie et al. [6] proposed a prediction method based on optimized GPR surrounding rock deformation and verified the predictive capability of the GPR prediction model in combination with actual engineering. Liu et al. [7] introduced Gaussian process regression (GPR) into tunnel construction surrounding rock deformation prediction to overcome the problems of existing models and won the comparison with support

vector machines. Zhang et al. [8] also improved and studied the application of GPR in the prediction of surrounding rock deformation and proposed a time series-based GPR prediction model for its application in predicting surrounding rock deformation during the tunnel construction period; in terms of neural networks, most of the research results for surrounding rock deformation prediction tend to favor the traditional neural network prediction model in the prediction of surrounding rock deformation in underground caverns. The ability of traditional neural network prediction models in predicting the surrounding rock deformation of underground caverns is inferior to that of SVM and GPR [9,10]. On this basis, in order to make the comparative work of this paper more meaningful, this paper introduces the long and short-term memory network (LSTM) for comparative study. In terms of research on LSTM, Yang et al. [11] proposed an LSTM-based method for concrete dam deformation prediction and verified the effectiveness of the model in concrete dam time series deformation prediction by example. Hu et al. [12] developed an LSTM-based combined prediction model for dam deformation and compared it with SVM in an example, arguing that LSTM has a more excellent prediction performance. Hu et al. [13] demonstrated the excellent performance of LSTM on time series prediction. All these studies demonstrate the potential of LSTM in predicting the deformation of surrounding rock in the underground cavern.

The above study demonstrates the feasibility as well as the potential capability of machine learning for surrounding rock deformation prediction. However, little work has been done to compare the application of these computational intelligence methods in the prediction of surrounding rock deformation, and the study of comparative work is of guidance for the application of these methods in surrounding rock deformation, which can comprehensively evaluate the prediction performance and prediction accuracy of various methods. In addition, for LSTM, its prediction work in the underground cavern surrounding rock deformation has hardly been carried out, so this paper also verifies whether LSTM can predict the tunnel surrounding rock deformation. The purpose of this research is to conduct a comparative study of computational intelligence methods, including gaussian process (GPR), support vector machine (SVM), and long short-term memory network (LSTM) in surrounding rock deformation prediction. These three techniques are typical computational intelligence methods that are widely used in engineering applications. In order to conduct a comparative study, the example of the Guzeng diversion tunnel in Liangshan, Sichuan Province, China, was taken. Divided into single-step prediction and multistep prediction modeling, these three computational intelligence methods are applied to the prediction of surrounding rock deformation in the tunnel, and the results of the three methods are analyzed and compared.

2. Methodology

2.1. Long Short-Term Memory (LSTM). Long short-term memory (LSTM) [14] is used as a special type of recurrent

neural network (RNN) in the autoregressive (AR) mode, which is deliberately designed to avoid long-term dependency problems and can be designed to remove the standard “healthy” components from the signal. Compared with RNN, LSTM can achieve long-term dependence and avoid the occurrence of gradient explosion and gradient disappearance. LSTM designs effective “gates” within the framework of each cycle to form different structures [15] that allow information to be selected for passage; that is, input gates, forgetting gates, and output gates are used to protect and control information. Since LSTM networks can learn long-term correlations in sequences, they do not require a prespecified time window and can accurately model complex multivariate sequences [16]. x_t and h_{t-1} are the input at moment t and the hidden layer at moment $t-1$, respectively. The three gates at the core of typical LSTM, the forgetting gate f_t , the input gate i_t , the output gate O_t and the information storage unit \tilde{C}_t that are updated to completion at moment t can be expressed as follows [17]:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f). \quad (1)$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i). \quad (2)$$

$$O_t = \sigma(W_O x_t + U_O h_{t-1} + b_O). \quad (3)$$

$$\tilde{C}_t = \tanh(W_C x_t + U_C h_{t-1} + b_C). \quad (4)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t. \quad (5)$$

$$h_t = O_t \odot \tanh(C_t). \quad (6)$$

These three gates control the forgetting of the long-term state C_{t-1} at the previous moment, the preservation of the temporary state \tilde{C}_t at the current moment, and the transfer of the long-term state C_t to the hidden layer h_t at the current moment, respectively. The gates are the forgetting gate, the input gate, the output gate, and the parameter matrix of the new information from the input layer to the hidden layer, respectively, where W_f , W_i , W_O , and W_C are the forgetting gate, the input gate, the output gate, and the parameter matrix of the new information from the input layer to the hidden layer, respectively. U_f , U_i , U_O , and U_C are the cyclic parameter matrix, and b_f , b_i , b_O , and b_C are the deviation matrix; σ is the sigmoid function, and \odot is the multiplication operator of the matrix. The long-term state is processed through the tanh layer, and then the result O_t of filtering through the output gate is multiplied to output h_t .

LSTM can achieve long-term dependence, and it can effectively avoid gradient explosion and gradient disappearance. Facing problems and tasks that are sensitive to time series, LSTM is usually more suitable. As a nonlinear model, LSTM with special implicit units can preserve the input for a long time, which can provide powerful guarantees for time series-based prediction. These features make LSTM good for the task of predicting the surrounding rock deformation in underground caverns.

2.2. SVM. The SVM method is now available for multi-variate classification problems and is a supervised learning model [18,19]. In addition to data classification, prediction and regression are extensions of support vector machines to mathematical problems and are known as support vector machine regression [20,21]. SVR models can generate regression functions by applying a set of high-dimensional linear and nonlinear functions using kernel functions. SVM applies the concept of ε -excitation loss function and is highly robust in prediction. Its core is to determine a separation hypersurface to make the desired risk as minimal as possible.

Under the given sample, $D = \{(x_i, y_i)\}_i^n$, where $\mathbf{x}_i \in \mathbb{R}^n$ is the input vector, and y_i is the target vector. Our final goal is to obtain the optimal $f(x) = \omega\phi(x) + b$ to approximate the target vector, with the deviation ε being strictly limited. Therefore, the parameters to be determined can be calculated by the following equation:

$$\min_{\omega, b} \frac{1}{2}\omega^2 + C \sum_{i=1}^n (\zeta_i + \zeta_i^*). \quad (7)$$

$$\begin{aligned} f(x_i) - y_i &\leq \varepsilon + \zeta_i, \\ y_i - f(x_i) &\leq \varepsilon + \zeta_i^*, \\ \zeta_i, \zeta_i^* &\geq 0, i = 1, \dots, n. \end{aligned} \quad (8)$$

The first term of the function is the canonical term, and the second term is the error risk. C is the penalty factor. To find the equations of the parameters to be determined, or the parameters ω and b to be determined in the objective function, the Lagrange function is constructed on this basis, the Lagrange multipliers are introduced, and the following constraints are added:

$$\begin{aligned} L(\omega, b, \alpha, \alpha^*, \zeta, \zeta^*, \mu, \mu^*) &= \frac{1}{2}\|\omega\|^2 + C \sum_{i=1}^n (\zeta_i + \zeta_i^*) - \sum_{i=1}^n \mu_i \zeta_i \\ &\quad - \sum_{i=1}^n \mu_i^* \zeta_i^* \\ &\quad + \sum_{i=1}^n \alpha_i (f(x_i) - y_i - \varepsilon - \zeta_i) \\ &\quad + \sum_{i=1}^n \alpha_i^* (f(x_i) - y_i - \varepsilon - \zeta_i^*). \end{aligned} \quad (9)$$

Converting to the pairwise problem of SVR, the solution of SVR is

$$f(x_i) = (\omega^T \phi(x_i) + b) = \sum_{i=1}^n (\alpha_i^* - \alpha_i) K(x_i, x) + b. \quad (10)$$

where α_i and α_i^* are for Lagrange multipliers and b is a scalar threshold. Clearly, the support vector is the only element of the data points used to determine the decision function. The above approach can be further extended to nonlinear surfaces, and SVR uses kernel function $K(x_i, x)$ to transform

the nonlinear inputs into a high-dimensional feature space without the need for realistic computational mapping.

SVR has been successfully applied to solve various time series prediction problems, and it performs very well in avoiding overfitting problems. For time series-based data such as surrounding rock deformation during underground cavern excavation, these characteristics of SVR are very suitable for prediction.

2.3. Gaussian Process Regression. Gaussian process is a stochastic process indexed by dimensions such as time and space. The essence of the Gaussian model is a regression model for uncertainty prediction of the distribution of output variables using a nonparametric probabilistic model [22,23] whose variables show the nature of a normal distribution for any linear combination; that is, the Gaussian process is a generalization of the Gaussian distribution. The form of GPR developed from GP is a variant of “inert learning” in machine learning. The Gaussian distribution reflects the distribution between random variables, while the Gaussian process represents the distribution between functions. The purpose of regression modeling is to establish a functional relationship $y = f(x)$ that can describe the training data. For the output variable y , it is assumed that the output values satisfy the following equation:

$$Y_i = f(X_i) + \varepsilon_i. \quad (11)$$

Here, $X = [X_1, X_2, \dots, X_N]$ denotes the input samples, $Y = [Y_1, Y_2, \dots, Y_N]$ denotes the observations, $f()$ denotes an arbitrary basis function, ε_i denotes Gaussian noise with zero mean, and N is the number of samples.

The observed value Y and the predicted value Y_* should satisfy

$$\begin{bmatrix} Y \\ Y_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} k + \sigma^2 I & k_{X_*} \\ k_{*X} & k(X_*, X_*) \end{bmatrix}\right). \quad (12)$$

$k = [k_{i,j}]$ and $k_{*X} = [k(X_*, X_1), \dots, k(X_*, X_N)]$, where $k_{i,j} = k(X_i, X_j) = \text{cov}(Y_i, Y_j)$

The predictive distribution $p(Y_*|Y)$ for Y_* should be a Gaussian distribution with variance and mean:

$$\bar{Y}_* = k_{*X} k^{-1} Y. \quad (13)$$

$$\sigma_{Y_*}^2 = k_{**} - k_{*X} k^{-1} K_{X_*}. \quad (14)$$

In the GPR model, the performance of the most commonly used covariance is crucial [24], and the covariance between individual dimensional features can be modeled by a Gaussian process kernel (kernel). The common types of kernels are Squared Exponential Kernel, Exponential Kernel, Matern 3/2, Matern 5/2, and Rational Quadratic Kernel.

The GPR model has the advantages of easy implementation, adaptive acquisition of hyperparameters, and probabilistic significance of the predicted output. It is suitable for dealing with nonlinear complex and complicated regression problems. Based on the above characteristics,

GPR has become a classic algorithm for dealing with deformation of surrounding rock in underground caverns.

2.4. Analysis of the Advantages and Disadvantages of the Algorithms. The advantage of LSTM is that it can achieve long-term dependence and effectively avoid gradient explosion and gradient disappearance. LSTM has high sensitivity to time-series data set, and its built-in hidden unit can preserve the input for a long time, which is not available in GPR and SVR. The disadvantage of LSTM is that it cannot handle parallel problems efficiently and will result in higher time cost when processing the problem due to its complex model structure, which is not a problem with SVR and GPR. Moreover, LSTM requires a lot of parameter settings in terms of network topology, initial values of weights, and thresholds and has a certain degree of “black box problem,” and there is no credible predictive equation, unlike GPR and SVR, which both have explicit forms of predictive equations. The advantages of the GPR model are that it is easy to implement, the hyperparameters are adaptively acquired, and the predicted output is probabilistically meaningful, it has a higher explanatory power compared to LSTM, and it is more convenient than SVR in terms of parameter setting. The disadvantage of GPR is that the prediction model usually performs well in interpolation prediction, while it performs worse than LSTM and SVR in extrapolation prediction due to the difficulty of data coverage. The advantages of SVR are its strict mathematical logic, high prediction accuracy for nonlinearities, good generalization ability on the test set, higher processing ability for small data set than LSTM and SVR, and higher performance in dealing with overfitting problems than LSTM and SVR. The disadvantage of SVR is that it is overly sensitive to parameters and kernel functions and has inferior processing power to LSTM and SVR when large fluctuations in the predicted data set occur. Since each algorithm has its own advantages and disadvantages, it is necessary to carry out targeted research to compare the performance of different algorithms in practical scenarios, taking into account the characteristics of the tunnel project itself and the actual engineering arithmetic examples.

2.5. Cross-Validation Algorithm. Cross-validation is often used in cases of over- and underfitting due to model complexity and is a method of chunking data samples to form smaller subsets on which to base the analysis. Cross-validation is a model validation technique used to evaluate the generalization ability of a statistical analysis model on independent data sets. In the data-intensive case, the data set is divided into two parts, that is, the training set and the test set, and then the average of error is obtained as the final evaluation to reflect the real ability of the model. In general, the commonly used cross-validation methods are random subsampling validation, K-fold cross-validation, and leave-one-out cross-validation, depending on the form of the cutoff. Although many studies have investigated leave-one-out cross-validation, from a computational perspective, k-fold cross-validation may be more common. For k-fold cross-validation, the data is randomly divided into k

approximately equal-sized packets $D^{(1)}, \dots, D^{(k)}$, and for each $D^{(i)}$, satisfying $n_i = |D^{(i)}|$, $\sum n_i = N$. One packet at a time is used as the test set, and the remaining k-1 packets are used as the training set for training. The accuracy of the training is discriminated based on the performance metric CV of the classifier.

$$CV_{N,K} = \sum_{i=1}^k \frac{n_i}{N} MSE_i. \quad (15)$$

2.6. Dropout. For the overfitting problem in deep learning, Hinton proposed the Dropout mechanism in 2012 [25], which is also a typical example of integrated models, and it is a good solution to the time-consuming problem of integrated models.

The dropout mechanism works in such a way that, during the training of a neural network, some neurons are randomly selected and temporarily hidden (discarded) for a layer of the network in one iteration, and then the training and optimization are performed again. In the next iteration, some neurons continue to be hidden randomly, and so on until the end of training. Since it is randomly discarded, each minibatch is training a different network.

During training, each neural unit is retained with probability p (dropout discard rate is $1 - p$); during the prediction phase (testing phase), each neural unit is present, and the weight parameter ω is multiplied by p . The output is $p\omega$. The schematic diagram is as follows.

The reason for the need to multiply P in the prediction phase is that the output of a neuron in the previous hidden layer before dropout is x , and the expected value after dropout during training is $E = px - (1 - p)0$; in the prediction phase, the layer neurons are always activated and in order to keep the same output expectation and to get the same result in the next layer, it is necessary to adjust $x \rightarrow px$, where p is the probability that the median value of the Bernoulli distribution (0-1 distribution) is 1.

3. Engineering Application and Comparison

3.1. Case Study. Guzeng Hydropower Station is located on the mainstream of the Muli River in Muli County, Liangshan Prefecture, Sichuan Province. It is the fifth step of the “One Reservoir and Six Levels” hydropower plan for the mainstream of the Muli River (Shangtongba-Abudi River section). The tunnel is located on the left bank of the Muli River. Water is drawn from about 0.4 km downstream of the mouth of the Xiaogou River to the Grade I terrace on the left bank of the Muli River, about 300m downstream of the mouth of Mannianjigang, building ground-level plants to generate electricity. From the inlet to the regulator well, the diversion line is about 11.06 km long. The water diversion tunnel area crosses four large-scale faults, such as Sawa Fault and Mannianjigang Fault. Small faults, extrusion fragmentation zones, and interlayer misalignment zones are randomly developed in the rock body, and fold phenomena such as intralayer kneading and flexure are also more developed, and groundwater is more active in the ditch section, fault zone,

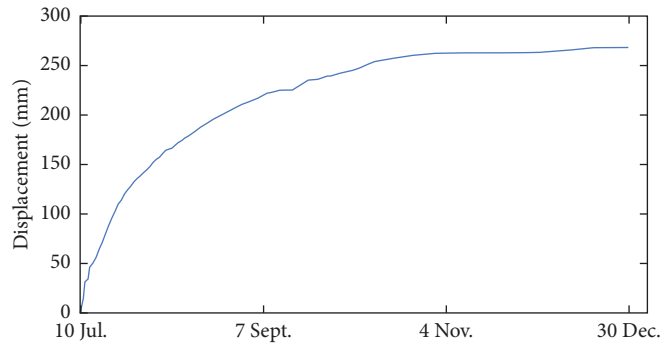


FIGURE 1: The surrounding rock displacement of Guzeng Tunnel.

and fissure-dense zone. Therefore, the condition of the tunnel surrounding rock is controlled by the rock strength, rock integrity, weathering, and unloading degree and also influenced by the spatial combination of rock layer, tectonic line and hole axis, groundwater, and other conditions. Rock strength and rock integrity are closely related to stratigraphic lithology; therefore, the principle of classification of surrounding rocks in tunnel area is based on lithology, combined with the development of geological formations, and the surrounding rocks in tunnel area are classified into III, IV, and V according to the geological classification standard of surrounding rocks in the Code for Geological Investigation of Small and Medium Hydropower Projects (DL/T5410-2009).

The data set used in this paper for the prediction of the surrounding rock deformation consists of the deformation data of the arch bottom surrounding rock of a typical section (section 7 + 318) of the main tunnel No. 4 of the Guzeng water diversion tunnel from July 10, 2019, to December 31, 2019, which was measured at the construction site by a multipoint displacement meter, as detailed in Figure 1 and Table 1. The hyperparameters of the LSTM, GPR, and SVR models are taken as shown in Table 2.

3.2. Analytical Results

3.2.1. Single-Step Prediction. Single-step prediction means that the input window predicts only one future value for each prediction. The strategy used in this paper for single-step prediction is to use the true value as an input to predict a future value and then input the true value in turn to obtain all future predicted values.

Figure 2 presents the results of SVR, GPR, and LSTM for single-step prediction of the surrounding rock displacement. As can be seen from the figure, all three methods can predict the amount of surrounding rock deformation in the latter 14 days well, and all of them can be good close to the actual data in terms of trend.

For the performance of single-step prediction models of SVR, LSTM, and GPR, we give data comparisons in terms of three evaluation metrics, MSE, MAE, and MAPE (Table 3). Comparing the data of the three evaluation metrics, we can see that the model effect of SVR is optimal compared with LSTM and GPR, while LSTM is slightly better than GPR, but

the model effect of both is not much different. Among them, the MAPE of SVR prediction model can reach 0.48%, which is much better than LSTM and GPR.

Moreover, from Figure 3, the prediction results of the three prediction models, SVR, LSTM, and GPR, all performed very well in the period from November 21 to December 31, almost fitting the original data.

3.2.2. Multistep Prediction. Multistep prediction means that each time a prediction is made, the input window predicts n future values (also called n steps). The strategy for multistep prediction in this paper is to use all true values as input when predicting the first future value, and when predicting the next $n-1$ values, the predicted values from the last prediction are used as input along with the true values for prediction, and one input is performed to obtain the future n values.

Figure 4 shows the effect of multistep prediction of surrounding rock displacement for SVR, GPR, and LSTM. From the figure, it can be seen that the three prediction models given in this paper all produce a certain degree of decline in the prediction effect of the surrounding rock deformation, but all three prediction models can still give a reference prediction effect on the surrounding rock deformation. Among them, the prediction data of SVR fits the data set more closely, while the prediction data of GPR is relatively close to the actual data in terms of trend but biased more in the prediction data from Oct. 6 to Oct. 31.

As for the performance of multistep prediction models of SVR, LSTM, and GPR, we still give data comparisons from three evaluation metrics, MSE, MAE, and MAPE, and the data are shown in Table 4. Comparing the data of the three multistep prediction model evaluation metrics, we can see that the model effect of SVR still outperforms that of LSTM and GPR, while the two models of LSTM and GPR are still closer in the two metrics of MAE and MAPE, but in the performance of MSE, GPR produced a larger error value of 7.12. Among them, the MAPE of SVR prediction model can reach 0.61%, and the prediction effect of the model is still excellent. And from Figure 5, the prediction effect of SVR multistep prediction model still performs very well very close to the data set during the period from November 21 to December 31, while LSTM and GPR are dwarfed.

TABLE 1: Surrounding rock deformation data.

Date	Deformation/mm
2019/7/10	0.000
2019/7/11	14.500
2019/7/12	31.370
2019/7/12	34.285
2019/7/13	46.175
2019/7/14	50.260
2019/7/15	55.880
2019/7/16	64.330
2019/7/17	71.455
2019/7/18	79.985
2019/7/19	88.355
2019/7/20	95.840
2019/7/21	102.580
2019/7/22	109.910
2019/7/23	114.030
2019/7/24	120.265
2019/7/25	124.495
2019/7/26	128.205
2019/7/27	132.62
2019/7/28	135.865
2019/7/29	138.600
2019/7/30	141.730
2019/7/31	144.540
2019/8/1	147.835
2019/8/2	151.990
2019/8/3	155.160
2019/8/4	157.240
2019/8/5	160.930
2019/8/6	164.335
2019/8/7	165.520
2019/8/8	166.530
2019/8/9	169.550
2019/8/10	172.260
2019/8/11	174.055
2019/8/12	176.655
2019/8/13	178.485
2019/8/14	180.655
2019/8/15	182.840
2019/8/17	187.745
2019/8/19	191.600
2019/8/21	195.715
2019/8/25	202.39
2019/8/28	207.475
2019/8/30	210.705
2019/9/1	213.070
2019/9/4	216.885
2019/9/7	222.100
2019/9/8	222.575
2019/9/11	225.055
2019/9/15	225.300
2019/9/20	235.305
2019/9/22	235.805
2019/9/23	236.090
2019/9/26	239.375
2019/9/27	239.390
2019/9/30	242.245
2019/10/4	245.155
2019/10/6	247.280
2019/10/9	251.560
2019/10/11	254.090

TABLE 1: Continued.

Date	Deformation/mm
2019/10/17	257.420
2019/10/23	260.300
2019/10/30	262.335
2019/11/8	262.805
2019/11/21	262.870
2019/11/28	263.070
2019/12/2	263.340
2019/12/12	265.791
2019/12/19	268.035
2019/12/24	268.140
2019/12/30	268.245

TABLE 2: Model hyperparameters.

Model	Hyperparameters
LSTM	input_size = 1; time_step = 6; batch_size = 1; num_layer = 2; hidden_size = 64; learn_rate = 0.002 drop_out = 0.2; optimization:Adam; Activation function:tanh
GPR	kernel = kernel1, n_restarts_optimizer = 100, alpha = 0.0029, random_state = 42
SVR	kernel = "rbf", C = 100, gamma = 0.1, epsilon = 0.0019

4. Discussion

Figure 6 and 7 show the performance of different surrounding rock deformation prediction methods in single-step prediction and multistep prediction. In recent years, scholars have been using LSTM to conduct research in various fields with a high degree of enthusiasm. Although LSTM has better processing performance for time-series tasks, its own model structure is relatively complex, and the time cost for training is much higher than that of SVR and GPR. Even though it alleviates the long-term dependency problem of RNNs and the "gradient disappearance" problem caused by the backpropagation of RNNs during training to a certain extent, it still cannot completely solve these problems. In addition, neural networks require a large number of parameter settings in terms of network topology, initial values of weights and thresholds, and all have a certain degree of "black box problem," and we cannot control the details of the processing of hidden nodes very well, and there is no credible prediction equation for neural networks. But GPR and SVR are different, and both of them have prediction equations of explicit form.

GPR is a nonparametric and kernel-based probabilistic model. The prediction model of GPR usually performs well in interpolation prediction, while for extrapolation prediction, GPR is not very good in extrapolation prediction due to the difficulty of data coverage. Compared with GPR, SVR requires more tuning parameters, but the processing of small sample data is the strength of SVM, so after using the appropriate method to optimize the parameters, SVR performs better in both single-step and multi-step prediction results.

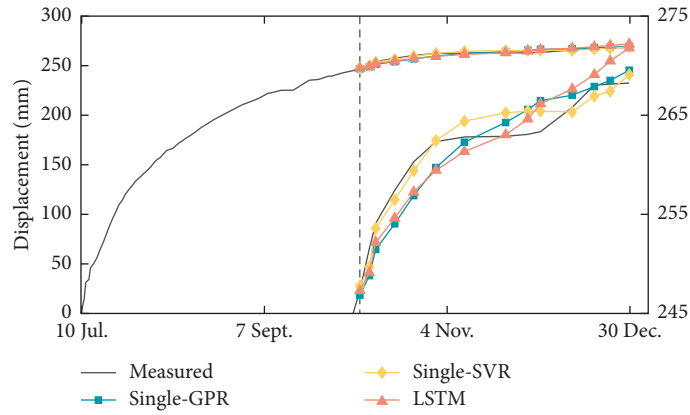


FIGURE 2: Prediction of surrounding rock displacement based on GPR, SVM, and LSTM (single-step prediction).

TABLE 3: Model evaluation index table (single-step prediction).

Model	MSE	MAE	MAPE
SVR	1.88	1.57	0.48
LSTM	4.67	1.89	0.74
GPR	5.09	1.99	0.77

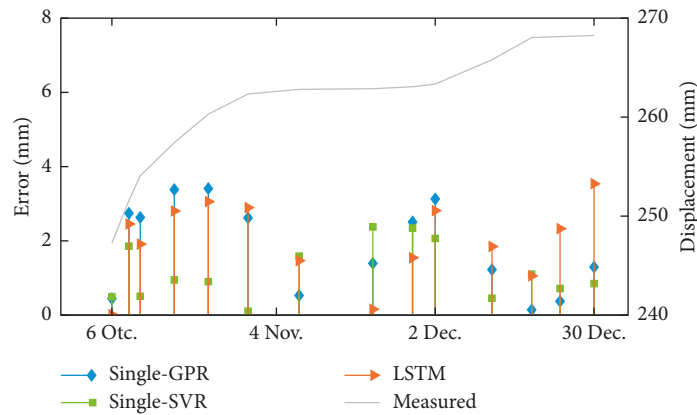


FIGURE 3: Displacement prediction and percentage error for different monitoring methods (single-step prediction).

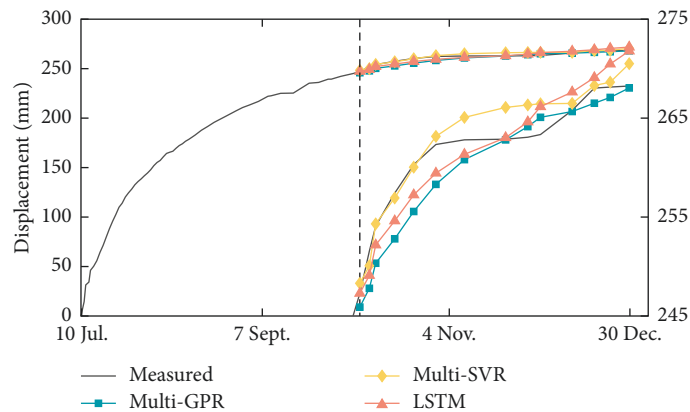


FIGURE 4: Prediction of surrounding rock displacement based on GPR, SVM, and LSTM (multistep prediction).

TABLE 4: Model evaluation index table (multistep prediction).

Model	MSE	MAE	MAPE
SVR	3.78	1.61	0.61
LSTM	5.26	2.09	0.78
GPR	7.12	2.08	0.81

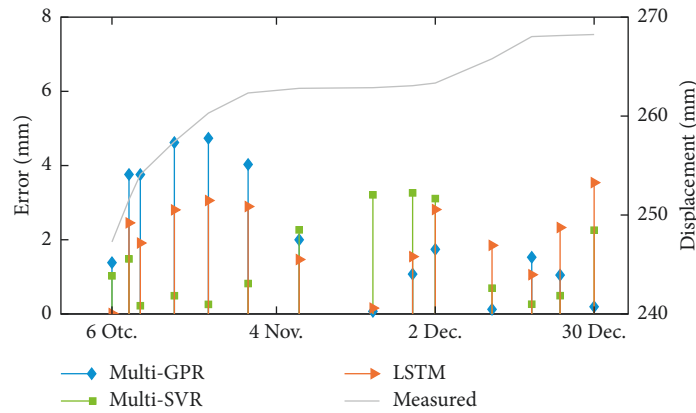


FIGURE 5: Displacement prediction and percentage error for different monitoring methods (multistep prediction).

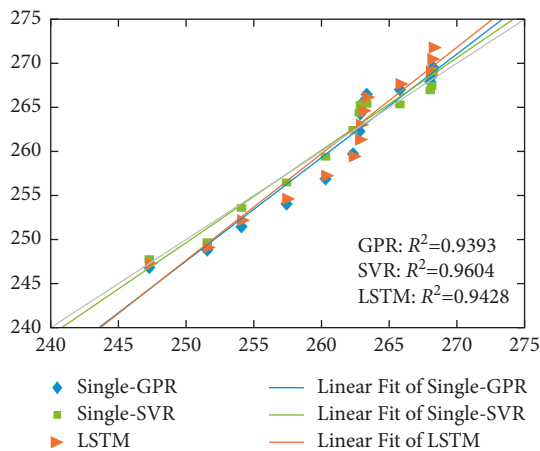


FIGURE 6: Performance of different methods for prediction of surrounding rock deformation (single-step prediction).

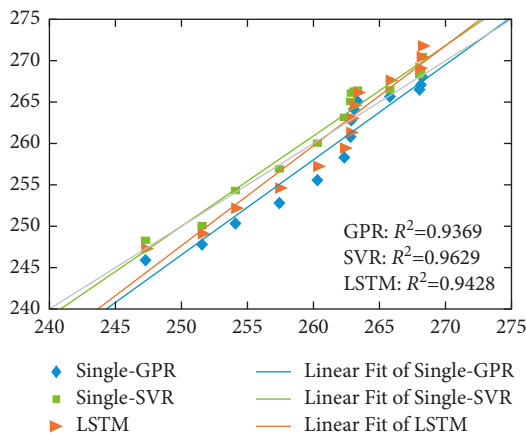


FIGURE 7: Performance of different methods for prediction of surrounding rock deformation (multistep prediction).

However, all prediction results depend on the quality of the data set and the ability to respond to the problem, and the setting of parameters can directly affect the performance of the prediction model. Therefore, the optimization and adjustment of various model parameters, as well as the analysis of the degree of influence of each parameter on the accuracy of the model, are the next tasks that can be attempted. In addition, if the confidence level and noise of the dataset used for model training is too high, the use of the trained model needs to be carefully considered in the context of the actual situation.

5. Conclusions

With the development of technology, the risk assessment and detection of underground cavern construction are getting more and more attention, and the prediction of surrounding rock deformation can provide a potential way to predict the large-scale deformation and collapse time of surrounding rock in the underground cavern, and in the future, the prediction of surrounding rock deformation is bound to be one of the important parts for the security of underground cavern construction and monitoring. The purpose of this paper is to compare the prediction effect of various algorithms for underground cavern surrounding rock deformation and provide some basis for the selection of algorithms for the prediction of surrounding rock deformation, and taking Guzeng diversion tunnel as an example, the single-step prediction model and multistep prediction model for surrounding rock deformation prediction are established by introducing LSTM, SVM, and GPR algorithms, and the prediction model performances of various algorithms are compared and analyzed. Firstly, from the prediction results of each model, it is feasible to use LSTM, SVM, and GPR algorithms to predict the deformation of the

underground cavern surrounding rock. In terms of performance, SVR outperforms LSTM and GPR in both single-step prediction and multistep prediction and is better than the other two in terms of predicting the direction of the surrounding rock deformation and the degree of data fit. In general, all these methods have better predictions for areas with stable deformation of the surrounding rock than for areas with large deformation of the surrounding rock. Therefore, the optimization of the performance of the intelligent algorithm model for the area with large fluctuations in the surrounding rock deformation is still a problem to be focused on for future research.

Data Availability

Some or all data, models, or codes that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge Yunnan Province Major Science and Technology Special Projects, China (Grant/Award Number: 202102AF080001).

References

- [1] A. T. C. Goh and S. H. Goh, "Support vector machines: their use in geotechnical engineering as illustrated using seismic liquefaction data," *Computers and Geotechnics*, vol. 34, no. 5, pp. 410–421, 2007.
- [2] B. Yao, C. Yang, B. Yu, F. Fang, and B. Yu, "Applying support vector machines to predict tunnel surrounding rock displacement," *Applied Mechanics and Materials*, vol. 29-32, pp. 1717–1721, 2010.
- [3] B. Yao, J. Yao, M. Zhang, and L. Yu, "Improved support vector machine regression in multi-step-ahead prediction for rock displacement surrounding a tunnel," *Scientia Iranica*, vol. 21, no. 4, pp. 1309–1316, 2014.
- [4] B. Yao, C. Yang, J. Yao, and J. Sun, "Tunnel surrounding rock displacement prediction using support vector machine," *International Journal of Computational Intelligence Systems*, vol. 3, no. 6, pp. 843–852, 2010.
- [5] S. Shi, R. Zhao, S. Li et al., "Intelligent prediction of surrounding rock deformation of shallow buried highway tunnel and its engineering application," *Tunnelling and Underground Space Technology*, vol. 90, pp. 1–11, 2019.
- [6] J. Xie and T. Lu, "A tunnel surrounding rock deformation prediction method with optimized Gaussian process regression," *Science Surveying and Mapping*, vol. 46, no. 04, pp. 50–56, 2021.
- [7] K. Liu, Y. Fang, and B. Liu, "Evolutionary Gaussian process regression model for tunnel surrounding rock deformation prediction," *Journal of the China Railway Society*, vol. 33, no. 12, pp. 101–106, 2011.
- [8] F. Zhang, "Research on tunnel deformation time series prediction based on multivariate GP-DE model," *Modern Tunneling Technology*, vol. 58, no. 01, pp. 109–116, 2021.
- [9] P. He, F. Xu, and S. Sun, "Nonlinear deformation prediction of tunnel surrounding rock with computational intelligence approaches," *Geomatics, natural hazards and risk*, vol. 11, no. 1, pp. 414–427, 2020.
- [10] Q. Wu, B. Yan, C. Zhang, L. Wang, G. Ning, and B. Yu, "Displacement Prediction of Tunnel Surrounding Rock: A Comparison of Support Vector Machine and Artificial Neural Network," *Mathematical Problems in Engineering*, vol. 2014, Article ID 351496, 6 pages, 2014.
- [11] D. Yang, C. Gu, Y. Zhu et al., "A concrete dam deformation prediction method based on LSTM with attention mechanism," *IEEE Access*, vol. 8, pp. 185177–185186, 2020.
- [12] A. Hu, T. Bao, and C. Yang, "A combined LSTM-Arima-based prediction model for dam deformation and its application," *Journal of Yangtze River Scientific Research Institute*, vol. 37, no. 10, pp. 64–68, 2020.
- [13] X. Hu, X. Sun, and L. Yin, "GPS coordinate time series prediction model based on multivariable LSTM," *Transducer and Microsystem Technologies*, vol. 40, no. 03, pp. 40–43, 2021.
- [14] X. Zhao, X. Han, W. Su, and Z. Yan, "Time Series Prediction Method Based on Convolutional Autoencoder and LSTM," in *Proceedings of the 2019 chinese automation congress (cac)*, pp. 5790–5793, Hangzhou, China, November 2019.
- [15] A. Graves, *Long Short-Term Memory*, pp. 1735–1780, Springer, Berlin Heidelberg, 2012.
- [16] P. Malhotra, L. Vig, G. Shroff, and P. Agarwal, "Long Short Term Memory Networks for Anomaly Detection in Time Series," in *Proceedings of the 23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, vol. 89, Belgium, 2015.
- [17] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM : a search space odyssey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2222–2232, 2015.
- [18] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer press, New York, NY, USA, 1995.
- [19] Y. Hu, "From pattern classification to active learning," *IEEE Signal Processing Magazine*, vol. 11, pp. 39–43, 1997.
- [20] T. K. James and W. T. Ivor, "Linear dependency between ϵ and the input noise in ϵ -support vector regression," *IEEE Transactions on Neural Networks*, vol. 14, no. 5, pp. 544–553, 2003.
- [21] A. J. Smola, N. Murata, and B. Schölkopf, "Asymptotically optimal choice of epsilon-loss for support vector machines," in *Proceedings of the Icann '98 Perspectives in Neural Computing*, Springer, London, 2008.
- [22] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine learning[M]*, The MIT Press, Massachusetts, 2006.
- [23] C. K. I. Williams, *Prediction with Gaussian processes: From Linear Regression to Linear Prediction and beyond[R]*, Aston University, Birmingham, 1997.
- [24] Z. Xiong, J. Zhang, and H. Shao, "Soft measurement modeling based on Gaussian processes," *Journal of System Simulation*, vol. 17, no. 4, pp. 793–795, 2005.
- [25] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," *Computer Science*, vol. 3, no. 4, pp. 212–223, 2012.