Int. J. Simul. Multidisci. Des. Optim. **15**, 24 (2024)

https://doi.org/10.1051/smdo/2024024

IJSMDO
Available online at:
https://www.ijsmdo.org

RESEARCH ARTICLE OPEN ACCESS

Security risk prediction technology for power monitoring system under the integration of OT and IT

Zhennan Zhu[1,*] and Jingquan Jin[2]

[1] School of Electronics and Information Engineering, Anhui Post and Telecommunication College, Hefei 230031, China
[2] School of Computer and Networking, Anhui Post and Telecommunication College, Hefei 230031, China

Received: 3 January 2024 / Accepted: 4 October 2024

Abstract. As an essential force for economic advancement and social stability, the security of the power system has always been a concern. Therefore, the security risks of power monitoring systems are a research focus. This study proposes a prediction method that integrates IT and OT for the security risk prediction of power monitoring systems. A basic indicator system for security risks for analyzing risk data is constructed, the support vector machine regression feature elimination method for predicting security risks in IT Technology is used. The experiment showed that the accuracy of the support vector machine regression feature elimination method was 92.35%, which was 6.06% higher than the error back propagation algorithm, 3.19% higher than the support vector machine algorithm, and 0.77% higher than the regression feature elimination algorithm. The maximum testing accuracy of the support vector machine regression feature elimination method was 0.96, which was 0.1 higher than the support vector machine algorithm, 0.04 higher than the regression feature elimination algorithm, and 0.17 higher than the back propagation algorithm. Therefore, the support vector machine regression feature elimination method can accurately predict power monitoring systems and has higher accuracy compared with other algorithms.

Keywords: Power monitoring system / safety risk prediction / support vector machine regression feature elimination method / IT technology / OT technology

1 Introduction

In daily life, electricity has become one of the most commonly used energy sources. The progress in various fields cannot be separated from electricity. Therefore, the secure implement determines the stable and secure advancement of society. The power monitoring system (PMS) is the core of the entire power grid system (PGS) and the guarantee for the normal implement. In the daily implementation of monitoring systems, they are threatened by external and internal devices. Ensuring the normal and stable implementation of PMS is currently a research focus, but many studies are unable to accurately predict PMS. IT technology is currently a widely used internet technology, and its fast computing speed and information processing speed have become the preferred solution to practical problems [1–3]. OT technology is a new operational technology that can handle real-world problems through the management and operation of IT technology. The traditional safety risk prediction methods for PMS have achieved risk prediction to a certain extent, but there are problems such as low prediction accuracy and poor stability. A new method based on the improved Support Vector Machine Regression Feature Elimination (SVM-RFE) algorithm is innovatively proposed for this research. The new method integrates IT technology and OT technology to construct a security risk prediction model for PMS. IT technology, with its powerful information processing and computing capabilities, provides rapid data analysis support for systems. OT technology ensures the security and reliability of system operations through real-time monitoring and operational management. The combination of the two not only enhances the security of the system, but also significantly improves the prediction accuracy and response speed [4]. The first part mainly introduces the research results all over the world. The second part parimarily concentrates on the construction of safety risk indicators and prediction systems for PMS. The third part is to conduct data analysis on the constructed model through experimental methods to study its predictive stability and accuracy. The fourth part summarizes the entire article content and provides prospects for future research directions.

* e-mail: zhuzhennan1983@outlook.com

2 Related works

Safety issues have always been a key focus of current power grid research. Many experts around the world have conducted extensive research on this. Among them, Wang Bo et al. found that the ship's power system (PS) may experience unstable electrical energy during operation, so a controller for voltage stability in the PS was proposed to stabilize the voltage. The new controller could learn and update extreme values in sequence data, which also achieved online prediction of PS. The experiment demonstrated that the new controller improved the stability of the PS while also improving the performance of wire voltage [5]. Shujun Chang et al. believed that traditional damping control methods were the key to solving the low-frequency oscillation problem in PS, but the delay in signal transmission and acquisition had a negative impact. Therefore, a new control method based on experimental prediction damping was proposed. The new method analyzed the stability of the PS and utilized derivative methods to determine the effectiveness of the model. The experiment demonstrated that the new method could markedly compensate for the delay signal of the PS, and possessed more advantages compared with traditional methods [6]. Tian Zhigang et al. found that many wind PS only analyzed the uncertainty of wind turbines and wind power, without considering the state issues of the PS. Therefore, a model for the reliability of wind power PS was proposed. The new model could meet the power requirements of wind PS. The experiment demonstrated that the new model could evaluate the PS state as well as improve the stability prediction of the system [7]. Hashemian, Seyed Mehran et al. found that excessive load may occur during the recovery period of the vehicle's PS, which also increased the demand for vehicle charging. Therefore, to avoid load issues, a method for simulating and predicting the charging demand of PS was proposed, which predicted the load during charging. The experiment indicated that the new method could accurately predict the PS load during electric vehicle charging [8].

Sobbouhi, Ali Reza et al. believed that the transient stability of the PS is the key to ensuring the stability prediction of the PS. Therefore, when studying traditional defects, a new method combining machine learning and time domain was used to classify the features of each data and set. The experiment indicated that the new method possessed more experimental accuracy relative to traditional power steady-state prediction [9]. Wei Cui et al. found that there are many power imbalances due to interference, which is very detrimental to the stability of the PGS. Therefore, to predict the frequency of the PGS and control the power grid state, a new confidence network model was proposed. This model could predict various information of the PGS. The experiment indicated that the new model had higher reliability and accuracy compared with traditional methods in the PGS [10]. Chengwei Fan et al. found in their frequency testing of PS safety states that using function derivatives was a key parameter for predicting PS frequency, making the prediction of safety states more stable. Therefore, a new algorithm was proposed. The experiment showcased that the new algorithm could predict the safety status of the PS [11]. Wang et al. believed that predicting the secure and stable state of the PS had great significance for the stability of the PS. However, traditional prediction methods cannot predict speed and accuracy very high. Therefore, a model training prediction method for PS safety steady-state was proposed, which predicted historical steady-state data. The experiment indicated that the new method improved the accuracy and performance of prediction [12].

In summary, currently, many predictions for PS are unstable and have low accuracy when conducting research. Meanwhile, these problems also exist in the safety risk prediction of PMS. Therefore, this study aims to address the low accuracy and insufficient stability of traditional prediction methods.

3 Establishment of a security risk prediction method for PMS integrating IT and OT

This chapter mainly analyzes the prediction indicators of safety risks in the PMS. IT and OT technologies are used to build the indicator system. Then, by integrating these two technologies, an algorithm for PMS safety risk prediction model is established.

3.1 Construction of basic security risk indicators for PMS integrating IT and OT

IT technology is a commonly used network information technology, usually referring to traditional information network technology. It covers a wide range, but the commonly used IT is mainly the application of some computer algorithms [13]. OT technology refers to network operation technology, which achieves the network operation by monitoring some coefficients and indicators. Therefore, monitoring and predicting indicators are particularly important in OT technology.

The key to the application of OT technology is to monitor and evaluate the predicted indicators in the safety risk prediction of PMS. The security risks predicted here mainly involve analyzing factors such as vulnerabilities and insecure network configurations in some networks. In the PMS, the operation of security systems includes terminal security issues related to network operations and security issues related to operations in the PS. Once there is a problem with the terminal system, it will bring incalculable security risks to the entire security system. Therefore, in the safety risk prediction of PMS, it is necessary to start with information technology. Strengthening information technology can achieve risk avoidance and improve security performance. The major components of the PMS are showcased in Figure 1.

In Figure 1, the PMS consists of a control hall, a management hall, and a data network. The control hall mainly manages and monitors the energy allocation and information transmission of the PMS. The management hall is mainly responsible for the automated management of information and the manipulation and management of trading systems. The data network mainly monitors and stores network data.

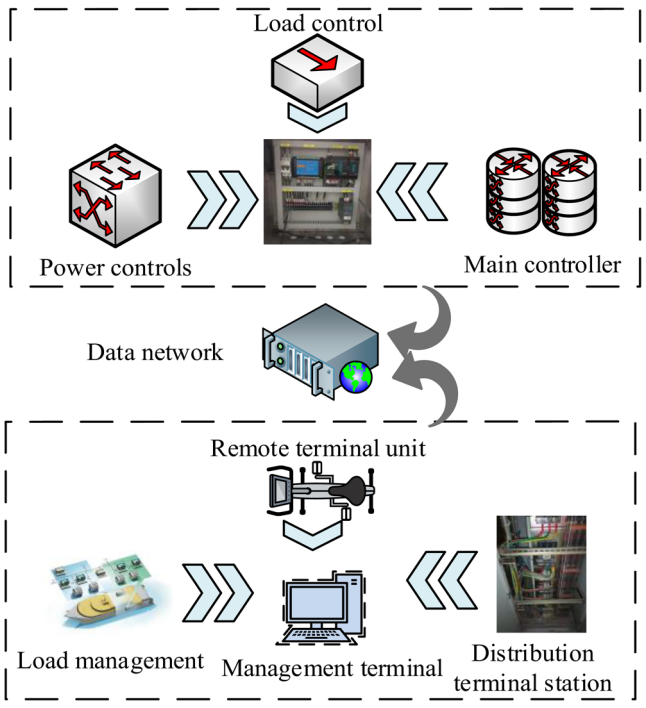

Fig. 1. Power monitoring system.

In the safety factors of PMS, not all safety risk indicators contribute to risk prediction, and many indicators have little effect on risk prediction. Therefore, to arrange the indicators in the prediction process, a new Support Vector Machine Recursive Feature Elimination (SVM-RFE) method is utilized. Meanwhile, to eliminate redundant relationships in the indicators of the PMS, Pearson is used to eliminate redundant relationships through correlation [14].

Currently, the main factors that cause risks in the PMS are separated into two types: external network attacks and internal personnel operations, both of which can lead to problems in the PMS. Therefore, risk prediction for PMS needs to start from these two aspects. However, the biggest threat to the PMS still comes from network attacks. Common network attacks include malicious intrusion, webpage attacks, scanning attacks, etc. Therefore, eliminating the complexity of indicators for network attacks is the most important risk prediction method for monitoring systems. Therefore, the SVM-RFE algorithm is used to classify data samples, as shown in equation (1), which is the feature dataset equation of the SVM-RFE algorithm.

$$T = \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_N \end{bmatrix} = \begin{bmatrix} x_{11} & \cdots & x_{1n} & y_1 \\ x_{21} & \cdots & x_{2n} & y_2 \\ \vdots & \vdots & \vdots & \vdots \\ x_{m1} & \cdots & x_{mm} & y_m \end{bmatrix}. \quad (1)$$

Equation (1), T represents the dataset of the sample. N represents the exponential dimension of the sample. x_m serves as the quantity of index values in the sample. y_m serves as the category label in the sample. Equation (2) is the decision function of the algorithm.

$$f(x) = sgn\left(\sum_{i=1}^{N} \alpha_i y_i x_i g x + b\right) \quad (2)$$

In equation (2), $f(x)$ represents the classification result obtained through the decision function variable. α represents the Lagrange factor. sgn represents the decision symbol of the function. b represents the displacement of the decision function. x represents the safety factor sample data of the PMS. The algorithm weight obtained through sample training is shown in equation (3) [15].

$$\begin{aligned} \omega &= \sum_{i=1}^{N} \alpha_i y_i x_i \\ &= \left(\sum_{i=1}^{m} \alpha_i y_i x_{i1}, \sum_{i=1}^{m} \alpha_i y_i x_{i2}, \cdots, \sum_{i=1}^{m} \alpha_i y_i x_{im}\right) \end{aligned} \quad (3)$$

In equation (3), ω represents the weight value of the element in question. It uses evaluation criteria to determine the forward and reverse selection samples of the algorithm, as shown in equation (4).

$$c_j = \omega_j^2. \quad (4)$$

In equation (4), c_j represents the evaluation score. When the score is higher, the correlation between feature extraction and classification coefficients will be stronger. ω_j represents the weight value at the jth indicator of the algorithm. The calculation for eliminating the redundancy of algorithm indicators is shown in equation (5).

$$pearson(x,y) = \frac{\sum_i (x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_i (x_i - \overline{x})^2}} \sqrt{\sum_i (y_i - \overline{y})^2}. \quad (5)$$

In equation (5), x and y represent two identical predictive indicators for the PMS. $\overline{x}$ and $\bar{y}$ represent the average values of the samples in the indicator. x_i represents the value of x indicator in the i-th sample. The value of the y indicator in the i-th sample for y_i increases the complexity of the entire prediction system as the correlation between the two indicators increases. For the selection of indicators for the security risk prediction system, redundant and miscellaneous indicators are first chosen, and then the remaining indicators are selected through feature set optimization. The selected feature set is mainly judged by the selection of classifiers. Another method is to judge the features of some indicators by judging the size of the feature set, as shown in equation (6).

$$g(S) = Max\left\{P(S) + \mu(1) - \frac{|S|}{|Q|}\right\}. \quad (6)$$

In equation (6), $g(S)$ represents a subset of features. $P(S)$ represents a subset of the sequence set Q that eliminates redundant indicators. $|S|$ serves as the quantity of indicators in the set S of indicators. $|Q|$ serves as the

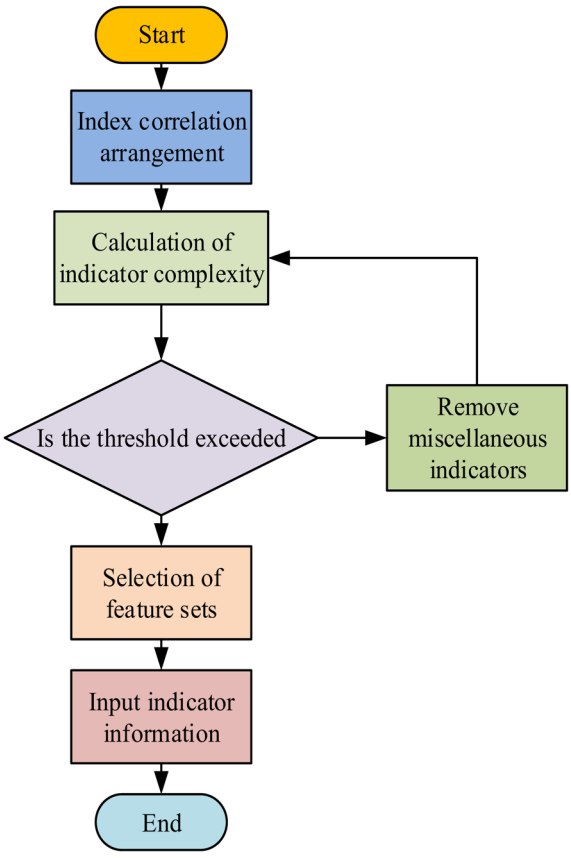

Fig. 2. Safety risk indicator system process.

quantity of indicators in the set Q of indicators. μ serves as the indicator weight factor for classifier classification. When the feature dataset value of the indicator is larger, the subset of the entire feature dataset is smaller. Therefore, the indicator construction is shown in Figure 2.

The safety risk indicator system in Figure 2 first arranges the correlation of the indicators, and then calculates the complexity of the indicators for determining whether they pass the threshold. If the redundancy is too high, the redundant indicators will be deleted. Otherwise, it will start selecting feature sets, input the selected feature sets into the indicator information system table, and finally end the indicator judgment.

3.2 Establishment of security risk prediction method for power monitoring system

In the construction of the PMS, IT and OT technologies can establish a PMS security risk indicator system. Pearson is then used for classifying the feature sets and eliminating redundancy of the indicators in the monitoring system safety risk prediction. Then, the SVM-RFE algorithm is used to classify and determine the remaining indicators. The obtained subdataset is analyzed and judged by the prediction system. The design is demonstrated in Figure 3.

The overall framework of the prediction system in Figure 3 starts with analysis and design. Firstly, the functional framework of the system is established. Then, the functionality of the software is designed through the system network structure. The functions of the system include data communication function, which mainly performs real-time communication detection of system data. The data analysis function mainly analyzes and solves data functions through IT technology. The status evaluation function mainly evaluates the status indicators of the current PMS. The monitoring and prediction function is to monitor and predict abnormal situations in the PS. The overall system function is capable of monitoring, managing, and storing data for use [16]. The PMS is different from the network security system, and its safety risk is composed of all components in the system. Therefore, when predicting security risks, models cannot be directly used for prediction. On the contrary, the electrical equipment used to predict safety risks in PMS should be evaluated first, and then the entire system should be predicted. The analysis of PMS is shown in Figure 4.

The overall safety risk prediction structure of the PMS, as shown in Figure 4, includes network devices, intelligent terminals, intelligent servers, databases, and other devices. Through the security risk prediction at the device layer, the security prediction of the monitoring system can be achieved. The SVM-RFE is a theoretical driven model that can process data based on the number of samples and has good generalization ability. Therefore, in predicting security risks, the regression linear used is shown in equation (7).

$$y = f(x) = w\phi(x) + b. \tag{7}$$

In equation (7), $\phi(x)$ serves as a nonlinear mapping in the spatial dimension. w represents the weight vector value. b represents the displacement of space in the function. The loss function of its function is calculated, as shown in equation (8).

$$L(f(x), y, \varepsilon) = \begin{cases} 0, |y - f(x)| \leq \varepsilon \\ |y - f(x)| - \varepsilon, |y - f(x)| > \varepsilon \end{cases}. \tag{8}$$

In equation (8), ε represents the maximum deviation in y and $f(x)$. Therefore, the computational problem of the algorithm can be represented by equation (9).

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n} L(f(x), y, \varepsilon). \tag{9}$$

In equation (9), C represents the penalty factor. When two relaxation factor equations ξ_i and ξ_i^* are added, equation (9) can be rewritten as shown in equation (10) [17]. Equation (10) defines the role of the objective function throughout the entire algorithm model, ensuring that the classifier can accurately classify data and prevent over-fitting.

$$\min\left(\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}(\xi_i, \xi_i^*)\right). \tag{10}$$

In equation (10), the Lagrange conversion factor is added to obtain the variation formula shown in equation (11). Equation (11) defines the kernel function and its

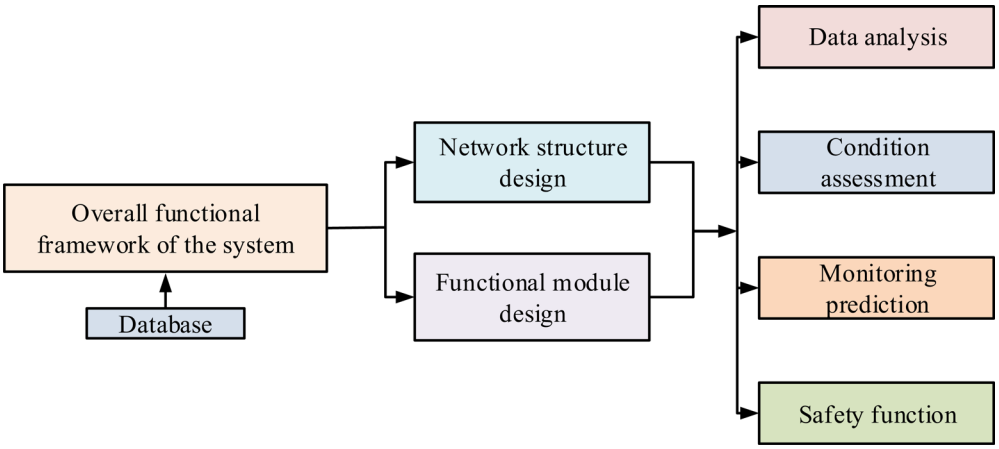

Fig. 3. System function design process.

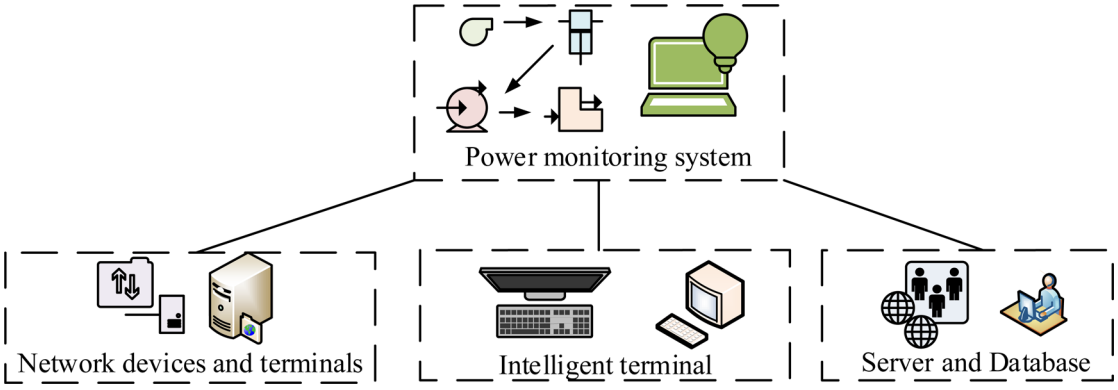

Fig. 4. Analysis of power monitoring system diagram.

parameters for mapping the input data to a high dimensional space for nonlinear classification and regression.

$$\max -\frac{1}{2}\sum_{i,j=1}^{n}(a_i - a_i^*)(a_j - a_j^*)K(x_i, x_j) - \sum_{i=1}^{n}a_i(\varepsilon - y_j) - \sum_{i=1}^{n}a_i^*(\varepsilon - y_i). \quad (11)$$

In equation (11), $K(x_i,x_j)$ serves as the Kernel Function (KF) with a value of $K(x_i,x_j) = \phi(x_i)\phi(x_j)$. a_i and a_i^* represent the Lagrange product factors. When using the equation to solve α and α^*, the equation for solving the shift and weight values is shown in equation (12). The model allows the constraints to be integrated into the optimization process by incorporating Lagrange multipliers, allowing the algorithm's constraints constraints to be satisfied during the optimization process, as shown in equation (12).

$$w = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\phi(x)$$

$$b = y_i + \varepsilon - \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)K(x_j, x_i). \quad (12)$$

In equation (12), due to the fact that the constraints of the equation cannot hold simultaneously, there is only one solution for α and α^*, and at least one solution exists to be 0. The final calculation of the algorithm is shown in equation (13). Equation (13) enables the model to make predictions on new data.

$$y = f(x) = w\phi(x) + b = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)K(x_i, x) + b. \quad (13)$$

In equation (13), $K(x_j,x_i)$ serves as the KF. The change in KF can transfer the entire algorithm to a higher dimensional space, thereby altering the linear nature of the problem. To achieve the smooth construction of the power monitoring safety risk prediction system, analyzing the system's functions is essential. Figure 5 is an analysis of the constructed system functional model [18].

In Figure 5, the system module functions include data processing and storage, as well as analysis and use of sample data. The data function module can also interact and

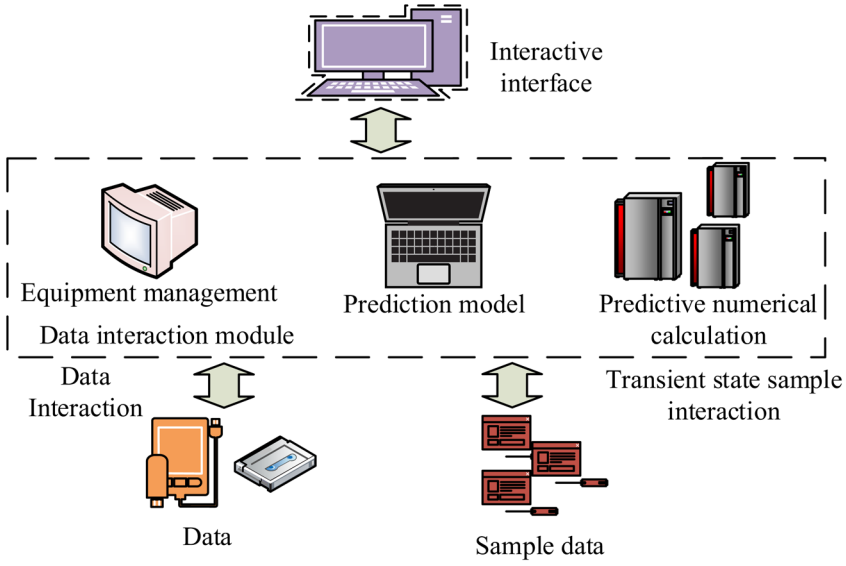

Fig. 5. System function module analysis.

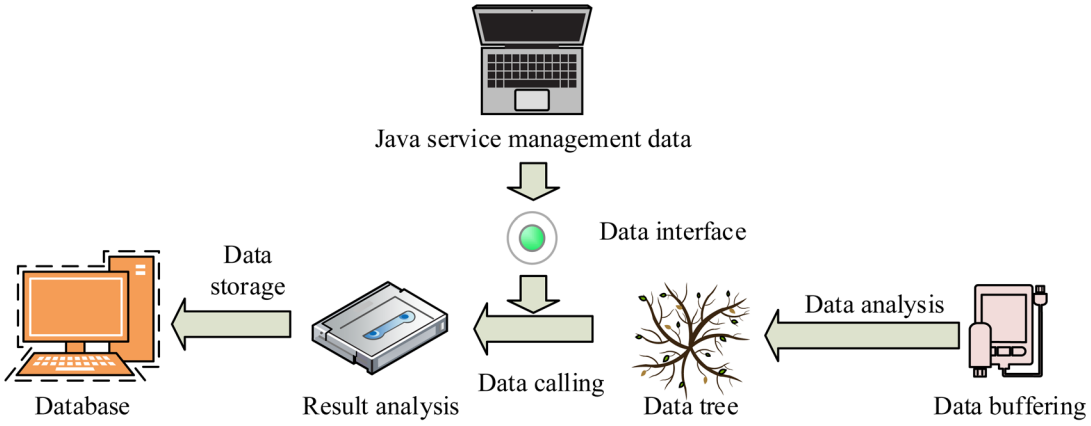

Fig. 6. Data analysis framework.

transmit user information data to achieve interoperability. When establishing a security risk prediction model for the PMS, the following two conditions need to be met. First, the equipment in PMS safety risks needs to predict safety risks, and then the safety risks of PMS need to be predicted through the predicted sequence values of PMS. Directly predicting the overall safety risks of PMS can lead to neglecting equipment status and affecting the overall prediction results. If all equipment cannot be predicted in the entire monitoring system, it will be impossible to predict the monitoring system from more details. Therefore, predicting devices can yield more complete prediction results. Due to the randomness of equipment safety risk prediction, it is necessary to extract linear features from the safety risk sequence during equipment prediction. Therefore, the safety risk prediction requires comprehensive evaluation of indicator data. Therefore, IT and OT techniques need to be used for analysis during the construction process [19].

The security risk prediction system model can collect data from security devices in chronological order, represented by the set function S. The S set contains multiple device security risk sequences and security prediction values at different times. In the prediction of security risks, if only time series changes are made to the security risks themselves, it will lead to inaccurate prediction results [20]. Therefore, the prediction requires data processing. Figure 6 displays the data analysis framework structure.

As shown in Figure 6, during the data parsing, data needs to be buffered, and the parsing results cause the data to be read. Then, data calls are made through the network tree of the data, using platform components for network calls. Finally, the data results are stored and imported into the database. After the data analysis is completed, it standardizes the data to ensure that each data has the same quantization structure and analysis effect, and then uses the SVM-RFE algorithm for processing the data. Meanwhile, for markedly enhancing the automatic learning in the predicted model, a gradient descent method is utilized for setting the learning efficiency control step size. Therefore, the commonly used method of automatic optimization, Adam optimization method, is added to the model, which eliminates the consideration of learning efficiency and can meet the autonomous learning needs of

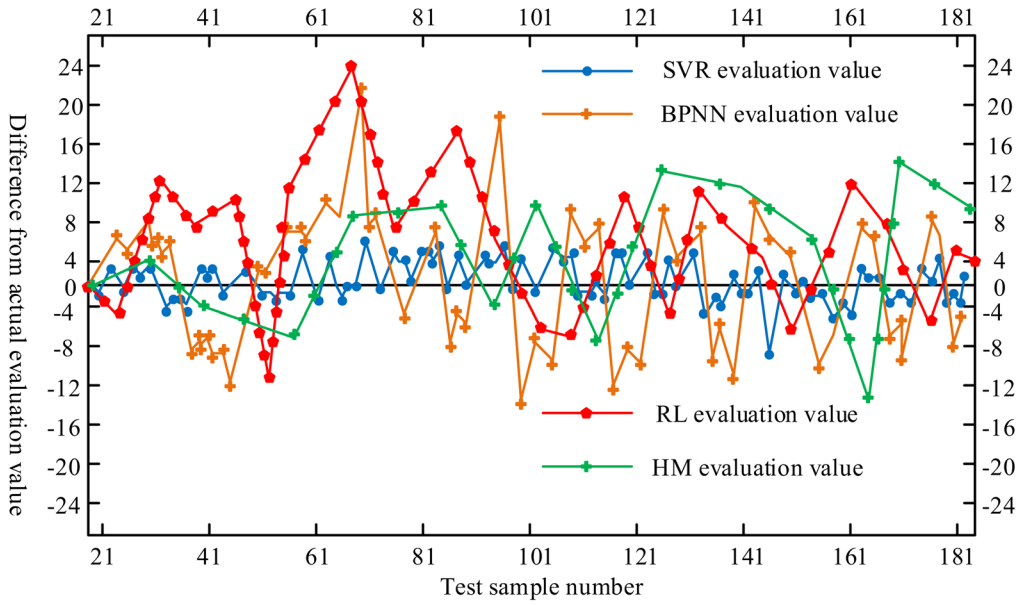

Fig. 7. Comparison of predicted values between two algorithms.

the model. This makes it possible to update and set up predictive models. In the process of power monitoring in the system, OT is used to monitor the real-time operation status of power equipment and collect key parameters such as voltage, current, temperature, and frequency. Simultaneously, IT system is used to collect network traffic logs, security event records, and user access behavior information [21]. Based on the Selective Support Vector Machine Regression Feature Elimination (SVM-RFE) algorithm, highly relevant features related to power system safety are extracted, including abnormal parameters of some power equipment, abrupt patterns of network traffic, and abnormal behavior of operators. Simultaneously, based on the extracted features, an SVM model is employed to construct a power system safety risk prediction model. During the model training process, the system optimizes the predictive ability of the model through repeated training and validation of historical data, enabling it to accurately identify potential risks [22]. After the algorithm model deployment is completed, the real-time data provided by the OT system are continuously input into the model for analysis. The system also predicts potential security risks in real time, and takes corresponding response measures based on the prediction results, issuing alerts, triggering security policies, or conducting preventive maintenance and a series of defense measures.

4 Analysis of experimental results on the safety risk prediction method for PMS integrated with IT and OT

Reinforcement learning (RL), hybrid models (HM), and traditional error back propagation (BP) models are used for power system prediction and analysis. 181 test datasets are used. The comparison of predicted values between the two algorithm datasets is shown in Figure 7.

As shown in Figure 7, the deviation between the predicted and actual values in the SVM-RFE algorithm fluctuates between 4 and −4, with several values having significant deviations exceeding 4. The maximum deviation value occurred when the sample size was 145, and the deviation value at this time was 9. The possible reason is that the sample data is too large, causing the algorithm to become unstable. The traditional BP algorithm had a sample fluctuation range between 12 and −12. There were several samples with significant deviations, with the maximum deviation occurring when the number of samples was 100 and the deviation value was −16. However, the BP algorithm had a large fluctuation in the prediction deviation value for safety risks, and many samples had significant deviations. The maximum deviation for the HM model occurred at sample 161, when the maximum deviation was 12. The predicted deviation for RL ranged from −12 to 24, where the maximum deviation occurred at 66 samples, with a maximum bias of 24. The designed algorithm is more stable, with smaller deviation values across the different sample models. To verify the predictive effect of the SVM-RFE algorithm on the prediction model, 90 experimental samples are selected and compared to obtain the prediction results, as shown in Figure 8.

As shown in Figure 8, the predicted value of safety risk increased with time when the time node was 33. The predicted value of the SVM-RFE algorithm in Figure 8a was close to the true curve change. In Figure 8b, the predicted value of security risk reached its peak and showed a downward trend. At this time, the predicted value of the SVM-RFE algorithm reached its peak simultaneously with the true value, and its predicted value was also the same. Figure 8c shows the changes when the security risk prediction decreases. The predicted values of the SVM-RFE algorithm had smaller deviations compared with other algorithms, making it more capable of predicting

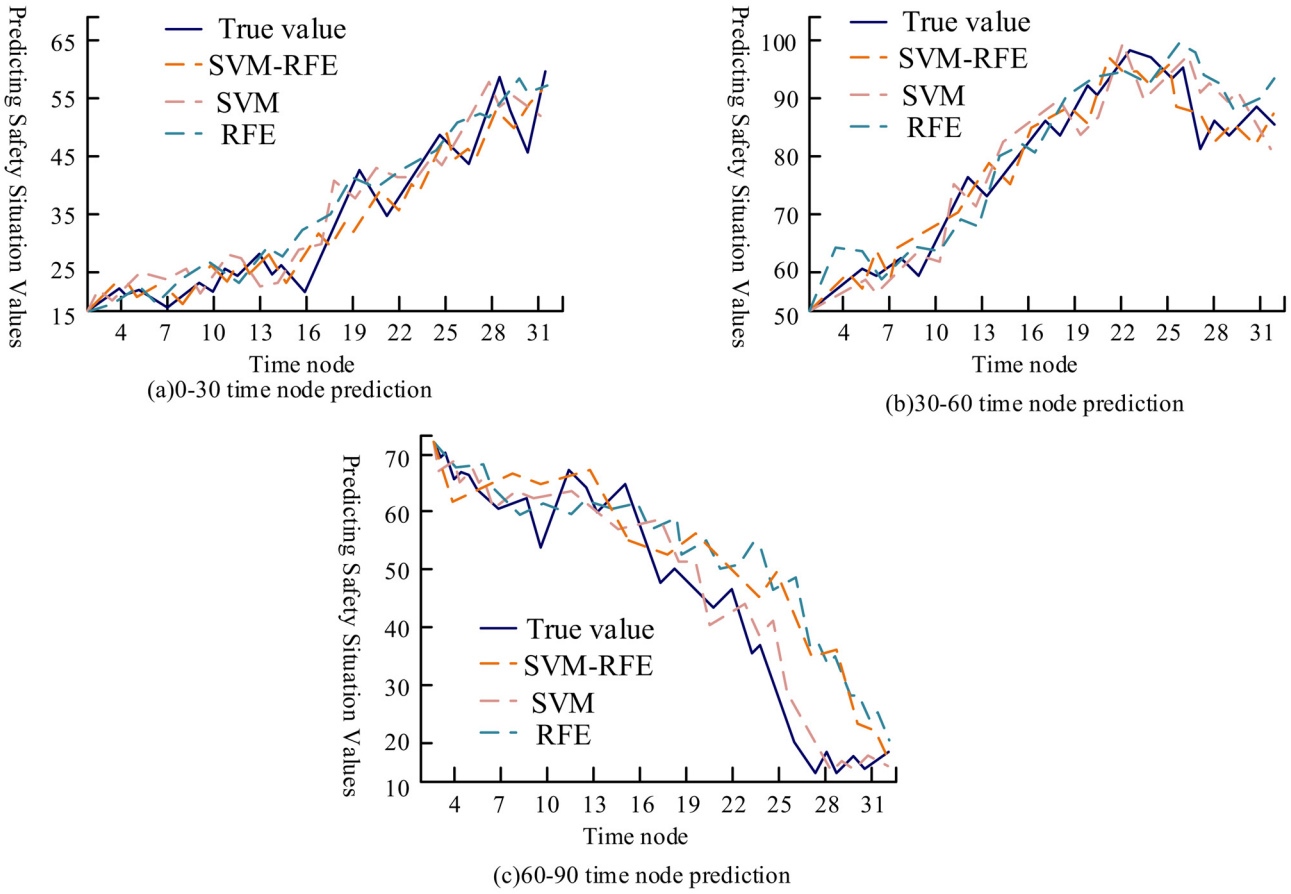

Fig. 8. Comparison of algorithm prediction effects.

security risks. Meanwhile, for testing the accuracy of the SVM-RFE algorithm, the accuracy is examined with other traditional algorithms, as shown in Figure 9.

As shown in Figure 9, the predicted value curve of the SVM-RFE algorithm among the four algorithms was similar to the actual curve trend, showing an overall fluctuating trend with an accuracy of 92.35%. The accuracy of the BP neural network was 86.29%, the accuracy of the Support Vector Machine (SVM) algorithm was 89.16%, and the accuracy of the Regression Feature Elimination (RFE) algorithm was 91.58%. This indicated that among the four algorithms, the SVM-RFE algorithm had higher accuracy, 6.06% higher than the BP neural network, 3.19% higher than the SVM algorithm, and 0.77% higher than the RFE algorithm. To verify the errors of the four algorithms under the same conditions, the average absolute error and root mean square error (RMSE) are compared to obtain the error curve changes, as shown in Figure 10.

As shown in Figure 10, the minimum trend of the average absolute error value of the SVM-RFE algorithm among the four algorithms ranged from 2.8 to 3.3, with a maximum error of 3.3, which occurred when the sample size was 90. The maximum error value of the BP algorithm ranged from 5.1 to 6.0, and its maximum error value occurred when the sample size was 90, with an error value of 6.0. The error values of SVM-RFE algorithm and BP algorithm were between 3.5 and 4.5. The RMSE and mean absolute error of the four algorithms have the same trend, but the difference lies in the size of the variation range. The RMSE of the SVM-RFE algorithm was between 3.0 and 4.0, with a maximum error of 4.0 when the sample size was 70. The error range of the BP algorithm was between 6.0 and 6.5, with a maximum error of 6.5 when the sample size was 300. The variation range of SVM algorithm and RFE algorithm was the same, ranging from 4.5 to 5.5, and the error trend of the two algorithms was not significantly different. To verify the accuracy of the SVM-RFE algorithm, the loss functions and accuracy are compared to obtain the variation diagram, as shown in Figure 11.

In Figure 11, the variation values of the loss functions of the four algorithms decreased as the iterations grew. Among them, the loss function of the SVM-RFE algorithm changed within 4 as the iteration was 15k. The loss function of the SVM algorithm changed within 5 when the number of iterations was 20 K. The loss function of the RFE algorithm was the same as that of the SVM algorithm, but the difference was that the range of the loss function changed as the iteration was 25 K. As the iterations of the BP algorithm were 30K, the loss function varied within 6. This demonstrates that the SVM-RFE algorithm has the lowest loss function, which is less than SVM algorithm and RFE algorithm 1, and less than BP algorithm 2. Therefore, the SVM-RFE algorithm is more stable. To verify the

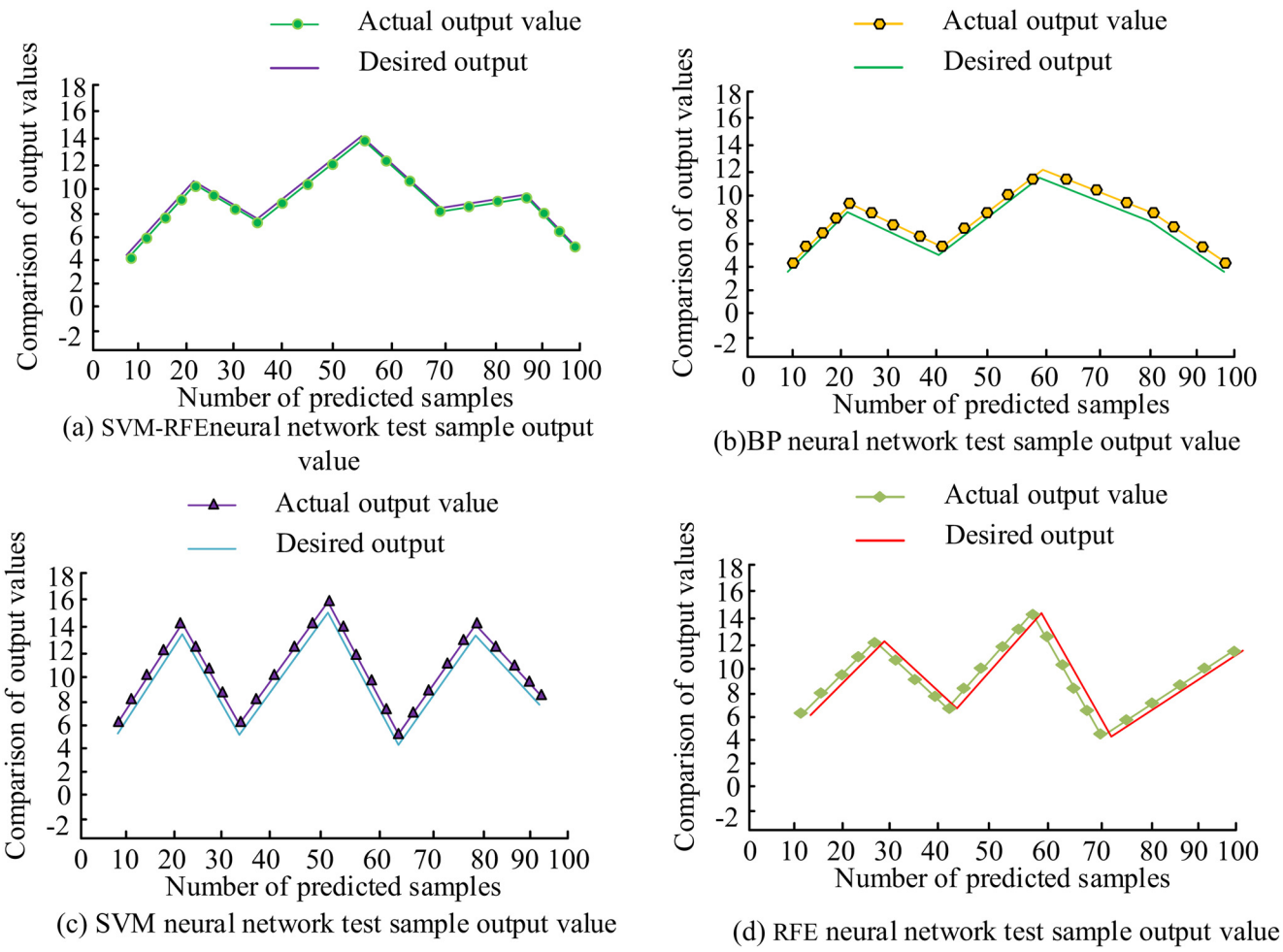

Fig. 9. Comparison of accuracy of four algorithms.

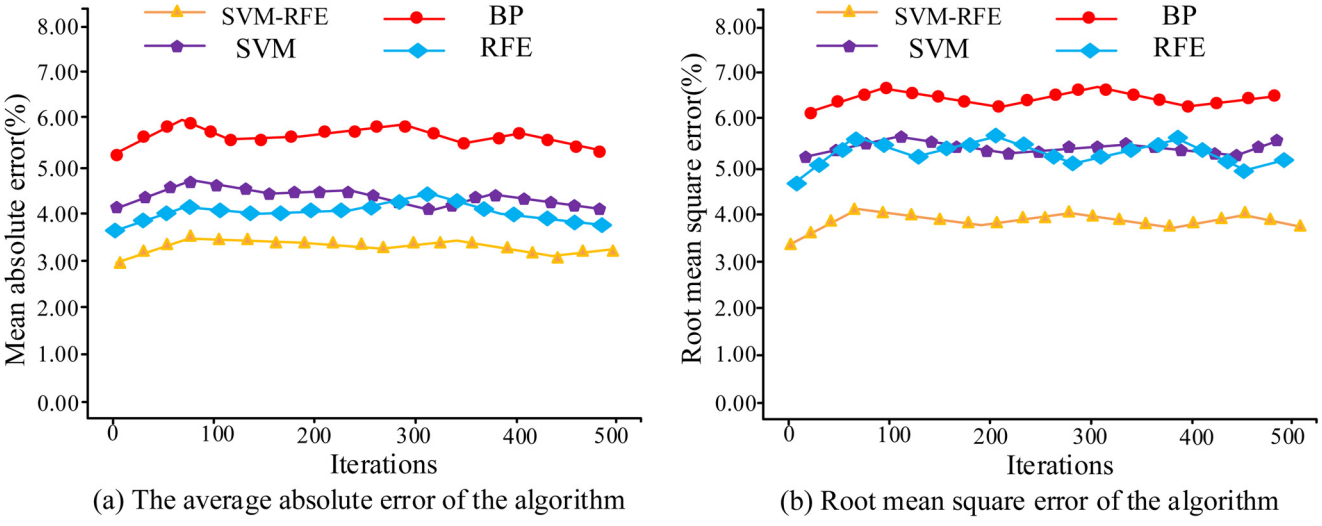

Fig. 10. Error variation curves of four algorithms.

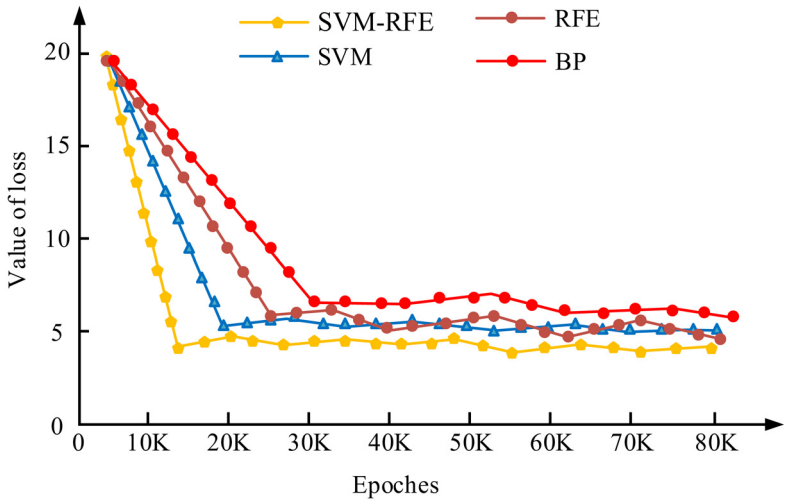

Fig. 11. Changes in loss functions of four algorithms.

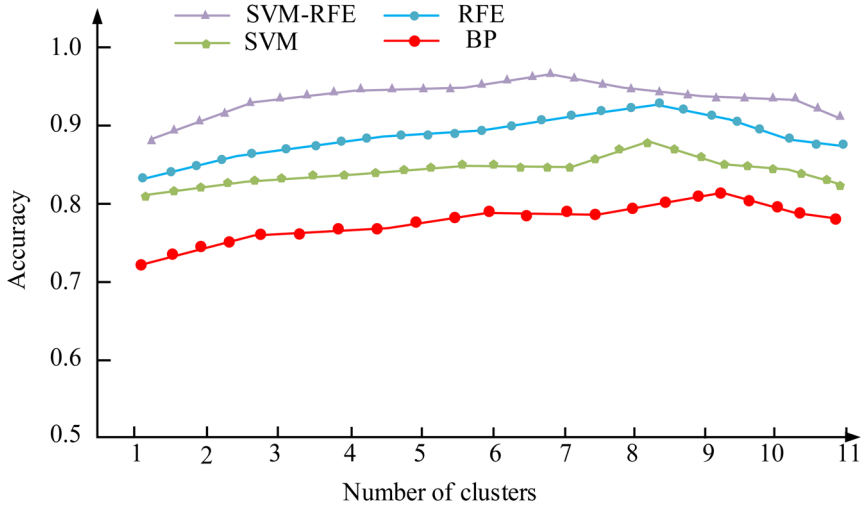

Fig. 12. Changes in frequency accuracy of four algorithms.

testing accuracy of the algorithm, four algorithms are compared to obtain the variation curve, as shown in Figure 12.

From Figure 12, the testing accuracy of the four algorithms showcased an increasing trend as the growth of sample size. The accuracy of the SVM-RFE algorithm was 0.96 when the sample size was 7, which reached its maximum value. The accuracy of the BP algorithm reached its maximum at a sample size of 9, with an accuracy of 0.79. The SVM algorithm reached a maximum value of 0.86 when the sample size was 8. Meanwhile, the RFE algorithm also reached a maximum value of 0.92 when the sample size was 8. This indicated that the testing accuracy of SVM-RFE algorithm was 0.1 higher than SVM algorithm, 0.04 higher than RFE algorithm, and 0.17 higher than BP algorithm. Among the four algorithms, the SVM-RFE algorithm is more suitable for predicting PS security risks. As shown in Figure 13, the flowchart compares the relationship among four algorithms, including SVM, RFE, SVM-RFE, and BP.

The SVM and RFE models in Figure 13 belong to the subordinate models of the SVM-RFE model. The SVM-RFE model incorporates SVM to classify and regress the selected features for data prediction, find the optimal decision boundary, and achieve prediction of safety risks. The RFE model is added to recursively eliminate all input features, evaluate the performance of the model after each elimination, gradually reduce the number of features, and ultimately retain the most important feature set. The BP model and SVM-RFE model are both data prediction and safety risk analysis models for PMS, and are in a parallel relationship.

5 Conclusion

This study utilized IT and OT technology to predict the security risks of PMS. Firstly, algorithm analysis was conducted on the indicator parameters of security risks, and a risk prediction indicator system was established

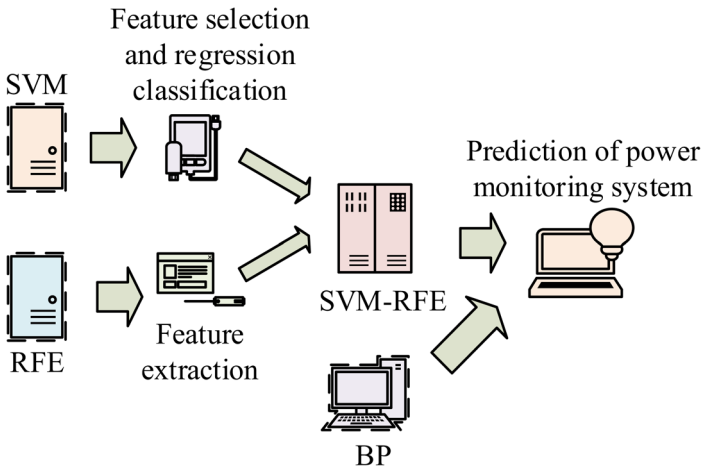

Fig. 13. Algorithm model relationship flowchart.

through the analysis of security risk indicators using OT technology. Then, IT technology and OT technology were used to construct a method for predicting security risks in the PMS, and a safety risk prediction model for the PMS on the ground of the SVM-RFE algorithm was established. The experiment showed that the accuracy of SVM-RFE algorithm was 92.35%, which is 6.06% higher than BP neural network, 3.19% higher than SVM algorithm, and 0.77% higher than RFE algorithm. The minimum trend of the average absolute error value of the SVM-RFE algorithm was between 2.8 and 3.3, with a maximum error of 3.3. The loss function of the SVM-RFE algorithm was the lowest, less than SVM algorithm and RFE algorithm 1, and less than BP algorithm 2. The test accuracy of the SVM-RFE algorithm was the highest at 0.96, higher than SVM algorithm 0.1, higher than RFE algorithm 0.04, and higher than BP algorithm 0.17. This indicates that the IT and OT technologies can predict security risks in PMS, while the SVM-RFE algorithm has better stability and accuracy in predicting security risks in PMS. Although this study has predicted security risks in PMS, there are still many shortcomings. First, in the analysis of safety risk indicators, many safety indicators have not been analyzed, so more indicators will be analyzed in the future. Second, the dataset and test sample size used are relatively small. More sample sizes will be tested in the future.

Funding

The research is supported by Anhui Province Key Natural Science Research Project (2022AH052959) Phase Research Achievements.

Conflicts of interest

The authors report there are no competing interests to declare.

Data availability statement

Data will be made available on reasonable request.

Author contribution statement

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Zhennan Zhu. The first draft of the manuscript was written by Jingquan Jin and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

References

1. Z. Lu, S. Lu, M. Xu, B. Cui, A robust stochastic stability analysis approach for power system considering wind speed prediction error based on Markov model, Comput. Stand. Interfaces **75**, 2–12 (2020)
2. J. Geng, X. Sun, F. Li, X. Wu, Prediction method of important nodes and transmission lines in power system transactive management, Electric Power Syst. Res. **208**, 10–18 (2022)
3. J. Shah, B. Mishra, IoT-enabled low power environment monitoring system for prediction of P M2. 5, Pervasive Mobile Comput. **67**, 10–26 (2020)
4. Y. Hu, H. Wang, Y. Zhang, W. Buying, Frequency prediction model combining ISFR model and LSTM network, Int. J. Electr. Power Energy Syst. **139**, 108–118 (2022)
5. B. Wang, X. Peng, L. Zhang, S. Peng, Adaptive bus-voltage control of ship power system with online fluctuation prediction, IET Generat. Trans. Distrib. **16**, 3109–3118 (2022)
6. S. Chang, C. Peng, Y. Hu et al., An improved prony prediction compensation-based wide-area damping control approach for power system low-frequency oscillation suppression, J. Sensors **2021**, 1–10 (2021)

7. Z. Tian, H. Wang, Wind power system reliability and maintenance optimization considering turbine and wind uncertainty, J. Qual. Mainten. Eng. **28**, 252–273 (2020)
8. M.R. Esmaili, A. Khodabakhshian, M. Gholipour, M.R. Esmaili, M. Malekpour, Approach for prediction of cold loads considering electric vehicles during power system restoration, IET Generat. Trans. Distrib. **14**, 5249–5260 (2020)
9. A.R. Sobbouhi, A. Vahedi, Transient stability prediction of power system; a review on methods, classification and considerations, Electric Power Syst. Res. **190**, 1068–1079 (2021)
10. W. Cui, W. Li, C. Wang, N. Yang, Y. Zhu, X. Bai, C. Xue, Prediction of primary frequency regulation capability of power system based on deep belief network, Int. Conf. Pioneering Comput. Sci. Eng. Educ. **20**, 423–435 (2020)
11. F. Chengwei, Y. Fei, X. Wang, Steady frequency prediction algorithm for power system with governor deadband, Eur. Trans. Electr. Power **28**, 1–14 (2018)
12. Q. Wang, C. Zhang, Y. Lü, Y. Zhihong, T. Yi, Data inheritance-based updating method and its application in transient frequency prediction for a power system, Int. Trans. Electr. Energy Syst. **29**, 2–17 (2019)
13. A.R. Sobbouhi, A. Vahedi, Transient stability improvement based on out-of-step prediction, Electric Power Syst. Res. **194**, 2–13 (2021)
14. Y. Fang, B. Luo, T. Zhao et al., ST-SIGMA: spatio-temporal semantics and interaction graph aggregation for multi-agent perception and trajectory forecasting, CAAI Trans. Intell. Technol. **7**, 744–757 (2022)
15. Z. Chen, Research on internet security situation awareness prediction technology based on improved RBF neural network algorithm, J. Comput. Cogn. Eng. **1**, 103–108 (2022)
16. A. Kazemian, Y. Basati, M. Khatibi, M. Tao, Performance prediction and optimization of a photovoltaic thermal system integrated with phase change material using response surface method, J. Cleaner Prod. **290**, 12–28 (2021)
17. G. Wang, Z. Zhang, Z. Bian, X. Zheng, A short-term voltage stability online prediction method based on graph convolutional networks and long short-term memory networks, Int. J. Electr. Power Energy Syst. **127**, 10–19 (2021)
18. J. Lin, J.A. Fernandez, R. Rayhana, F. Mengge, Predictive analytics for building power demand: day-ahead forecasting and anomaly prediction, Energy Build **15**, 255–272 (2022)
19. X. Liu, H. Wang, Q. Wang, G. Tongyao, Research on fault scenario prediction and resilience enhancement strategy of active distribution network under ice disaster, Int. J. Electr. Power Energy Syst. **135**, 10–27 (2022)
20. Y. Wang, X. Qi, Y. Chen, Enhanced fault localization in multi-terminal transmission lines using novel machine learning, Int. J. Simul. Multidisci. Des. Optim. **15**, 15 (2024)
21. S.D. Milić, Z. Đurović, M.D. Stojanović, Data science and machine learning in the IIoT concepts of power plants, Int. J. Electr. Power Energy Syst. **145**, 108711–108712 (2023)
22. V. Govindaraj, P. Palpandian, V. Varun, M. Ram, Automated power factor correction and predictive energy monitoring using machine learning, in *2024 International Conference on Science Technology Engineering and Management (ICSTEM)* (2024), pp. 1–7

Cite this article as: Zhennan Zhu, Jingquan Jin, Security risk prediction technology for power monitoring system under the integration of OT and IT, Int. J. Simul. Multidisci. Des. Optim. **15**, 24. (2024)