Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2025 SCPE. Volume 26, Issues 3, pp. 1377–1385, DOI 10.12694/scpe.v26i3.3900

DESIGN OF ELECTRICAL LOAD PREDICTION SYSTEM BASED ON DEEP LEARNING ALGORITHM

YUANLI XU*

Abstract. In order to solve the problems of low prediction accuracy and long model time in current prediction algorithms, the author proposes the design of an electrical load prediction system based on deep learning algorithms. The author proposes an improved deep learning short-term load forecasting model based on random forest algorithm and rough set theory. The model is first based on historical data and uses the random forest algorithm to extract key feature quantities that affect load forecasting. Then, the key feature quantities and historical load values are used as input and output terms for training the deep neural network, and the prediction results are corrected through rough set theory. Finally, simulation verification is conducted through numerical examples. The experimental results showed that compared with the RF-DL model, the MSE index of the RFDL-RST model decreased by 30.187%, and the overall prediction results were closer to the true values. The MAE index also decreased from 5.76% to 4.02%. During special periods of significant load changes such as 07:00-08:00 (rapid increase in load) and 22:00-23:00 (rapid decrease in load), the prediction accuracy was greatly improved. In addition, compared with the DL-RST model, the MAE and MSE indicators of the RF-DL-RST model were reduced by 15.210% and 21.414%, respectively, and the DL training time of the RF-DL-RST model was shortened by 10.175%, indicating that simplifying the DL input feature quantity through the RF model can improve the load forecasting effect. The prediction accuracy of this model is higher than that of a single deep learning model and a model without prediction correction.

Key words: Electrical load forecasting, Random Forest (RF) algorithm, Deep learning (DL), Rough Set Theory (RST)

1. Introduction. Predicting electricity demand in the near future is crucial for maintaining the safety and efficiency of power systems[1]. In recent years, with the continuous development of the global electrical market, the spot market and intraday trading system have been continuously improved, and the requirements for load forecasting accuracy have become increasingly high [2]. There are various factors that affect load, including weather factors (temperature, humidity, sunlight intensity, etc.) and time factors (working days, holidays, current specific time, etc.). At the same time, some policy factors can also lead to changes in load patterns, such as factory production reduction and shutdown caused by epidemic control, resulting in a decrease in electricity load; Encouraging policies for electric vehicles have led to an increase in electricity demand. The above factors make short-term load forecasting exhibit strong non-linear and stochastic characteristics [3]. Electrical load forecasting is a series of forecasting work that takes electrical loads as objects, including predicting future electrical demand (power), predicting future electricity consumption (energy), and predicting load curves [4]. Forecasting the load on electrical systems plays a vital role in planning and operating power grids, forming the basis for tasks like dispatching, regulation, and control. Typically, load forecasting spans various timeframes, including long-term, medium-term, short-term, and ultra short-term, each serving specific purposes within the electrical system [5,6]. Short-term load forecasting, extensively studied and highly relevant to experts and scholars, focuses on predicting electrical demand for the upcoming day to week. Typically, this involves forecasting the capacity or daily and weekly consumption data of a specific region. The aim is to establish power generation plans and inform operational scheduling. Various methodologies exist for short-term load forecasting, with intelligent forecasting methods currently being the predominant approach. This method applies powerful intelligent algorithms to establish a prediction model and make predictions. As the power grid continues to expand, its complexity grows, demanding more sophisticated load forecasting techniques. Intelligent algorithms like neural networks and support vector machines are commonly employed for this purpose. As the grid becomes smarter, there's a greater need for load forecasting methods that can meet the evolving demands of this dynamic environment. Traditional intelligent algorithms belong to shallow structure algorithms, and shallow structures

*Weifang Vocational College, Weifang, Shandong, 262700, China (Corresponding author, wzyjdxu@163.com)

are difficult to effectively represent nonlinear complex functions when given limited samples [7].

2. Literature Review. With the rapid progress of computer technology, machine learning is undergoing a resurgence, finding applications across diverse fields like image recognition, object detection, and natural language processing. In electrical load forecasting, machine learning algorithms have demonstrated considerable success. Advanced techniques such as reinforcement learning and transfer learning have been leveraged in this domain. However, traditional machine learning approaches face challenges with high-dimensional data. Deep learning, particularly Artificial Neural Networks (ANNs), addresses this by extracting meaningful features from complex data, thereby enhancing forecasting accuracy. ANNs, consisting of input, output, and hidden layers, serve as foundational tools in machine learning, facilitating the modeling of intricate relationships within the data. Sasidharan, M. P. et al. conducted an analysis of several prominent deep learning architectures, including Long Short Term Memory (LSTM), Gated Recurrent Unit (GRU), Convolutional Neural Network (CNN), as well as hybrid models combining LSTM with CNN, and CNN with GRU, among others. They explored the suitability of these models for predicting charging load in charging stations. Training data were sourced from multiple charging stations within a specific region to train the models. The performance of each model was evaluated using standard metrics, and the resulting predictions were thoroughly assessed and presented [8]. Jin, X. B. et al. introduced an attention-based encoder-decoder network enhanced by Bayesian optimization to achieve precise short-term power load forecasting. This model adopts an encoder-decoder structure incorporating a Gated Recurrent Unit (GRU) recurrent neural network, renowned for its robustness in handling time series data [9]. Zhu, C. et al. introduced a power load forecasting approach utilizing Long Short-Term Memory (LSTM) networks. Through comparative analysis with conventional models, their method demonstrated reduced errors and increased applicability, showcasing its effectiveness in power load prediction [10].

The author proposes an improved deep learning (DL) short-term load forecasting model (RF-DL-RST) based on RF algorithm and rough set theory (RST). This model introduces policy factors and, together with time and weather factors, establishes a load forecasting feature set. Key feature quantities and historical load values are used as inputs and outputs for deep learning training, and the prediction results are corrected through rough set theory. Simulate and verify the effectiveness of the model.

3. Method.

3.1. Introduction to Random Forest Algorithm. The schematic diagram of the random forest algorithm is shown in Figure 1. The key to the random forest algorithm lies in the decision tree, which obtains prediction or regression results by voting or weighted averaging the prediction results of each decision tree [11].

Researchers both domestically and internationally have developed numerous decision tree algorithms, including ID3, C4.5, and Classification and Regression Tree (CART). These algorithms employ a top-down methodology to construct decision trees [12]. In the process of forming a decision tree, each new node needs to choose a new attribute as the basis for splitting. The difference between these three decision tree algorithms lies in the decision criteria for leaf splitting during the growth process. Among them, CART uses minimum mean square error as the attribute metric for splitting regression trees and Gini index (GI) as the splitting criterion for classification trees. When using the random forest algorithm for classification, the final result is determined by voting. When using the random forest algorithm for regression, the prediction result is obtained by taking the mean [13,14]. In addition, in order to reduce the impact of overfitting and random errors on the prediction results, the original data is generally divided into training and testing sets, and then the bootstrap method is used to extract the training set. Then, the CART algorithm is used to train each decision tree from top to bottom one by one until it meets the requirements.

3.2. Feature extraction. How to select key feature quantities in the dataset is crucial for reducing model complexity and shortening computation time. When extracting key feature quantities using the random forest algorithm, the Gini index or out of bag data error rate is generally used for evaluation.

The author used the Gini index method for research, and the principle is as follows: Assuming the dataset

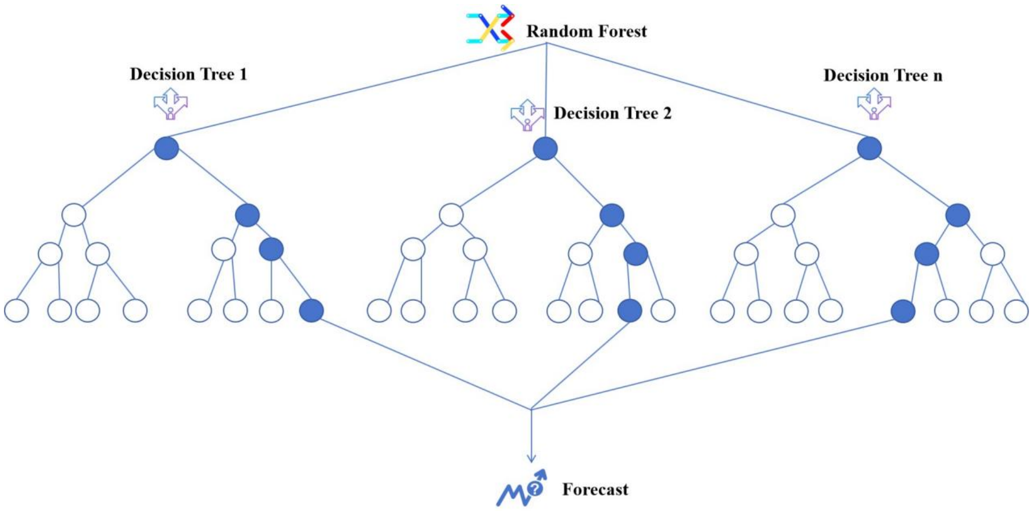

Fig. 3.1: Schematic diagram of the random forest algorithm

has J feature quantities $(X_1, X_2, X_3, \cdots, X_j)$, C categories, and I decision trees, the Gini index of node m is

$$G_m = \sum_{c=1}^{C} \widehat{p_{mc}}(1 - \widehat{p_{mc}}) \tag{3.1}$$

In the formula, $\widehat{p_{mc}}$ represents the probability estimate that node m sample is of class c.

The importance score V_{jm}^{GI} of feature quantity X_j at node m is represented by the change in Gini index before and after node m branching:

$$V_{jm}^{GI} = G_m - G_l - G_r \tag{3.2}$$

In the formula, G_l and G_r are the Gini indices of two new nodes 1 and r after node m branches.

If the set of nodes where the feature quantity X_j appears in the i-th tree is set to M, then the importance of the feature quantity X_j in the i-th tree is represented as

$$V_{ij}^{GI} = \sum_{m \in M} V_{jm}^{GI} \tag{3.3}$$

In summary, the importance of the feature quantity X_j in RF can be expressed as

$$V_j^{GI} = \frac{1}{I} \sum_{i=1}^{I} V_{ij}^{GI} \tag{3.4}$$

From this, it is possible to rank the importance of each feature quantity in the dataset and extract important feature quantities.

3.3. Principles of Deep Learning. A Deep Neural Network (DNN) is a type of neural network architecture that consists of multiple layers, including at least one hidden layer [15]. Compared with traditional BP neural networks, the two have similar structures, but DNN generally has more hidden layers and adopts a layer wise training mechanism to overcome the gradient diffusion problem in BP neural network training. Compared with traditional solving methods, well trained DNNs have higher computational efficiency and accuracy.

A typical DNN network structure consists of input and output layers at the beginning and end, with the middle layer being the hidden layer and the layers being fully connected (any node in the previous layer must be connected to any node in the following layer). Assuming there are g nodes in the i-1st layer, the output h_j^i of the jth node in the i-th layer is represented as

$$h_j^i = \sigma(z_j^i) = \sigma(\sum_{k=1}^{g} \omega_{jk}^i h_k^{i-1} + b_j^i) \tag{3.5}$$

In the formula: $\sigma(\cdot)$ is an activation function used to sum and further enhance the input of a node; ω_{jk}^i is the weight coefficient from the k-th node in layer i-1 to the j-th node in layer i; h_k^{i-1} is the output of the k-th node in the i-1 layer; b_j^i is the deviation coefficient of the jth node in the i-th layer.

The author uses the mean squared error loss function, represented as follows:

$$L = \frac{1}{PT} \sum_{p=1}^{T} \sum_{t=1}^{T} (y_{p,t} - p'_{pt})^2 \tag{3.6}$$

In the formula: P represents the number of training samples; $y_{p,t}$ is the expected value of the p-sample at time t; $y'_{p,t}$ is the predicted value output by DNN; T is the number of predicted time periods.

At the same time, the author introduces L2 regularization to the loss function, aiming to limit the weight parameters to a certain range to adapt to outliers and noise. The expression is as follows:

$$L = \frac{1}{PT} \sum_{p=1}^{p} \sum_{t=1}^{T} (y_{p,t} - y'_{p,t})^2 + \frac{\alpha}{2} \omega^T \omega \tag{3.7}$$

In the formula: α to regularize hyperparameters; ω for weight vectors.

Set the learning rate of the parameter as μ, update the hidden layer parameters repeatedly through equation 3.7 until the prediction accuracy converges.

3.4. Predictive Correction Model. Rough set theory is a mathematical tool for dealing with uncertainty and fuzzy problems, which can effectively correct and analyze defect information that is inconsistent, requires error correction, or has data loss [16].

Establishing a load forecasting correction model using rough set theory:

$$\begin{cases} y'_{t+1} = y_{t+1} + s_t |k_{t+1} - k_t| \\ k_{t+1} = y_{t+2} - y_{t+1} \\ k_t = y_{t+1} - y_t \end{cases} \tag{3.8}$$

In the formula, y_{t+1} and y'_{t+1} are the predicted and corrected values at time t+1, respectively; s_t is the scale factor.

In order to solve the scaling factor s_t, an information system needs to be constructed. The author assumes that the information system on which the rough set theory is based is K=(U,A), where: the domain U is the set of predicted values output by DNN; $A = C \cup S$ is the attribute set, S=st represents the decision attribute, and the conditional attribute C is the set of feature quantities in the dataset. Based on existing research results, C={a,b,c} is defined here [17]. Among them:

$$a = \frac{|k_{t+1} - k_t|}{y_t} \tag{3.9}$$

$$b = sgn(k_{t+1} - k_t) \tag{3.10}$$

$$c = |\frac{y_t}{max(y_t)}| \tag{3.11}$$

At this point, the load prediction value can be corrected through equations 3.9-3.12.

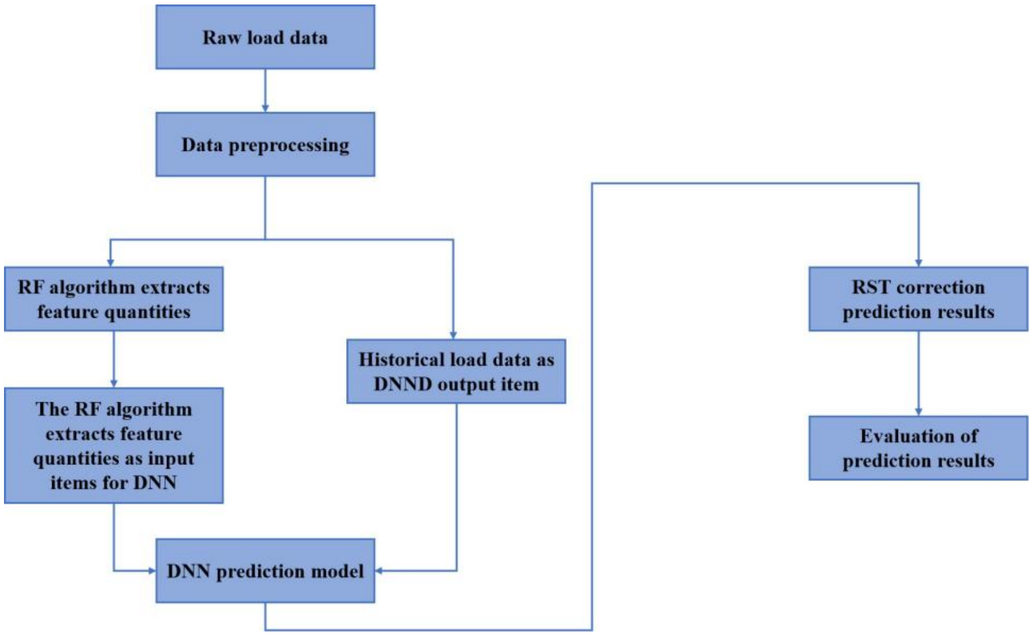

Fig. 3.2: Schematic diagram of RF-DL-RST model

3.5. Prediction Result Evaluation Model. The author employs two metrics, mean square error (MSE) and maximum absolute error (MAE), to assess the accuracy of the prediction results. MSE quantifies the overall prediction performance by measuring the average squared difference between predicted and actual loads. On the other hand, MAE evaluates the predictive accuracy at specific points by calculating the maximum absolute difference between predicted and actual loads. The MSE and MAE are indicated as follows:

$$\epsilon_{MSE} = \frac{1}{N}\sum_{n=1}^{N}(y_n - y_n')^2 \tag{3.12}$$

$$\epsilon_{MAE} = max(|\frac{y_n - y_n'}{y_n}|) \tag{3.13}$$

In the formula, N represents the number of predicted points; y_n is the true value of the nth predicted point; y_n' is the predicted value of the nth prediction point [18,19].

3.6. RF-DL-RST prediction model. The RF-DL-RST model framework is shown in Figure 3.2. The author's goal is to make short-term predictions of electrical loads, and the input feature quantities include various factors such as weather and time, which differ from the predicted results (that is load data) in terms of dimensions, units, etc. Therefore, preprocessing of the predicted data is necessary[20].

There are many factors that affect the electricity load in a region, including weather, time, and policies. However, the prediction accuracy of DNN is not positively correlated with the input items. When there are too many input items, it not only causes the network structure to be complex, but also may degrade the model accuracy.

The author establishes a feature set for load forecasting. However, the author believes that the week date, workdays, and holidays in its time factor constitute duplicates, so the characteristic quantity of the week date is excluded. At the same time, considering the impact of epidemic lockdowns on social electricity consumption in recent years, the author also studied whether the day was under lockdown as a characteristic quantity. In addition, the author also added weather factors such as average temperature, average wind speed, sunrise time, and sunset time as characteristic variables. The specific predicted feature quantities are shown in Table 3.1.

Table 3.1: Predicted feature quantities

Influence factor	Feature	Meaning
Time factor	Month	From January to December
	Day	Specific dates of each month
	Weekday	Normal work, value 1
	Festival and holiday	Saturdays, Sundays, and other holidays, with a value of 0
	Day Hour	00:00-24:00
Weather factors	Maximum temperature	The highest temperature of the day, ℃
	Minimum temperature	The lowest temperature of the day, ℃
	Average temperature	Average temperature of the day, ℃
	Average relative humidity	Daily average humidity,%
	Weather conditions	Such as sunny, cloudy, rainy, snowy, etc
	Air quality	Air quality index
	Average wind speed	Daily average wind speed, m/s
	Sunrise time	Specific time
	Sunset time	Specific time
Policy factors	Is it under lockdown or not	When affected by epidemics or natural disasters, take 1, otherwise take 0

3.7. Experimental Analysis. The author used load data from a certain regional power grid from October 28, 2022 to February 4, 2023 to simulate and verify the RF-DL-RST prediction model. In order to verify the superiority of the RF-DL-RST model, two comparative models are set up, among which: Comparative model 1 is the RF-DL model without RST correction part; Comparison Model 2 is a DL-RST model without RF feature selection. The selection of relevant parameters for the three models is consistent.

4. Results and Discussion.

4.1. Key feature extraction for load forecasting. Rank the importance of the predicted feature quantities selected in Table 3.1. The Random Forest (RF) model is configured with 500 decision trees and utilizes 3 split features. The dataset is divided into training and testing sets at a ratio of 9:1. Figure 4.1 illustrates the analysis findings regarding the importance of various features.

From Figure 4.1, it can be seen that after sorting the 15 feature quantities in Table 1 in order of importance scores from low to high, the 8 feature quantities of the day, including hours, minimum temperature, average temperature, weather conditions, holidays, workdays, sunrise time, and whether they are under control, have higher scores. Therefore, they are used as input items for the DNN model.

4.2. Deep learning training. Train the DNN model using the 8 key feature quantities and historical load data selected by RF as input and output items, respectively. The number of input layer nodes in DNN is 8, and the number of output layer nodes is 1. Set DNN to have 3 hidden layers with 40, 30, and 20 nodes respectively, and activate ReLU function; The ratio of training set to test set is 9:1, and the training frequency is 200 times. During the iteration process, the mean square error of the predicted values varies with the number of training iterations, as shown in Figure 4.2. It can be seen that the mean square error begins to converge at around 150 training iterations and continuously approaches the value of 975MW^2.

4.3. RST correction. According to equations 3.8-3.12, calculate the conditional attributes C={a,b,c}, as well as the decision attribute S before t, in order to obtain the rough set information system. Given the requirements of rough set theory for processing data, the encoding rule for the conditional attribute C={a,b,c} is set as follows:

$$C = \{a \in [1, 6], b \in [1, 3], c \in [1, 6] | a, b, c \in Z\} \tag{4.1}$$

From this, the corrected load forecasting data can be calculated.

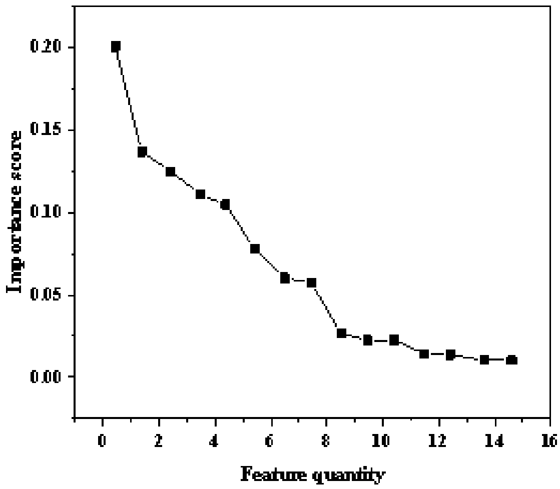

Fig. 4.1: Analysis results of feature importance in random forest algorithm

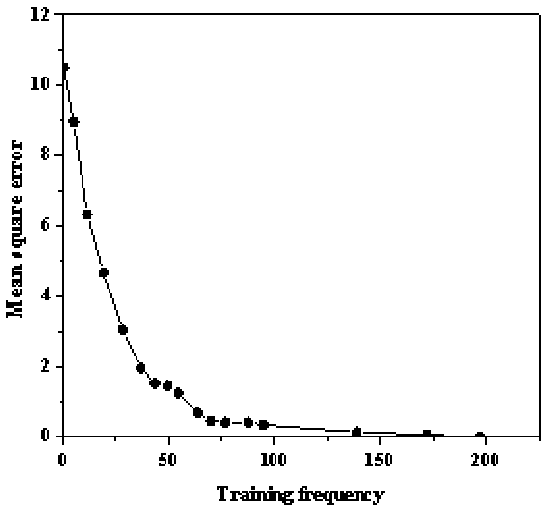

Fig. 4.2: Curve of Mean Square Error of Predictions as a Function of Training Times

Figure 4.3 shows the actual load on February 5, 2023 and the predicted load curve before and after RST correction. It can be seen that the predicted load curve after RST correction is basically between the actual load curve and the predicted load curve without RST correction, and is closer to the actual load curve.

4.4. Comparative analysis. According to equations 3.13 and 3.14, the evaluation indicators for the predicted results can be calculated. The comparison of indicators between RF-DL-RST model and RF-DL, DL-RST models is shown in Table 4.1.

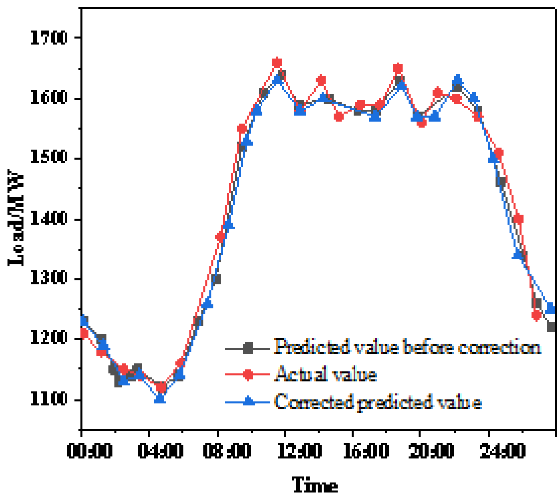

Fig. 4.3: Comparison of actual load and predicted load curve before and after RST correction

Table 4.1: Comparison of Indicators for Three Models

Model	DL training time/s	MSE MW^2	MAE/%
RF-DL-RST	96.28	680.32	4.02
RF-DL	96.28	974.64	5.76
DL-RST	107.20	865.83	4.72

From Table 4.1, it can be seen that compared with the RF-DL model, the MSE index of the RFDL-RST model has decreased by 30.187%, and the overall prediction results are closer to the true values. The MAE index has also decreased from 5.76% to 4.02%. During special periods of significant load changes such as 07:00-08:00 (rapid increase in load) and 22:00-23:00 (rapid decrease in load), the prediction accuracy has greatly improved. In addition, compared with the DL-RST model, the MAE and MSE indicators of the RF-DL-RST model were reduced by 15.210% and 21.414%, respectively, and the DL training time of the RF-DL-RST model was shortened by 10.175%, indicating that simplifying the DL input feature quantity through the RF model can improve the load forecasting effect. Based on the above analysis, it can be concluded that the RF-DL-RST model has significantly better prediction results, which verifies the effectiveness of the author's prediction model.

5. Conclusion. The author suggests a design for an electrical load forecasting system leveraging deep learning algorithms. Specifically, for short-term load forecasting, they introduce the RF-DL-RST model, which combines the random forest algorithm with rough set theory. Through example calculation and analysis, the following conclusions can be drawn: 1) By evaluating the importance of factors affecting load through RF, the model calculation time is shortened and the accuracy of prediction is improved. 2) By modifying the model results through RST and establishing evaluation models from both global and local perspectives, the effectiveness of the method was verified, greatly improving the accuracy of predicting load sudden changes.

6. Acknowledgement. Institute of Education Science, Chinese Academy of Management Science Research on the Cultivation of Students' Logical Reasoning from the Perspective of Core Literacy. ZGYJKS4090KT

REFERENCES

[1] Zhang, R., Yu, M., & Zhang, C. (2022). A similar day based short term load forecasting method using wavelet transform and lstm. IEEJ Transactions on Electrical and Electronic Engineering, 17(4), 506-513.
[2] Hao, L., Linghua, Z., Cheng, T., & Chenyang, Z. (2023). Short-term load forecasting model based on gated recurrent unit and multi-head attention, 30(3), 25-31.
[3] Wang, H., Zhang, N., Du, E., Yan, J., Han, S., & Liu, Y. (2022). A comprehensive review for wind,solar,and electrical load forecasting methods, 5(1), 22.
[4] Agyemang, F., Yamoah, S., & Debrah, S. K. (2022). Pseudocritical rapid energy dissipation analysis of base-load electrical demand reduction on nuclear steam supply system, 12(2), 19.
[5] Liu Bairu. (2022). Integration,coordination and empowerment-view the keywords of the"14th five-year plan"on photovoltaic development from the perspective of china's modern energy system planning, 29(2), 41-45.
[6] LI Xiao, & LU Xianling. (2022). Ethod for forecasting short-term power load based on dual-stage attention mechanism and gated recurrent unit network. Computer Engineering, 48(2), 291-296,305.
[7] Xu, L. H., Huang, C. Z., Wang, Z., Liu, H. L., Huang, S. Q., & Wang, J. (2024). Novel intelligent reasoning system for tool wear prediction and parameter optimization in intelligent milling, 12(1), 76-93.
[8] Sasidharan, M. P., Kinattingal, S., & Simon, S. P. (2023). Comparative analysis of deep learning models for electric vehicle charging load forecasting. Journal of The Institution of Engineers (India), Series B. Electrical eingineering, electronics and telecommunication engineering, computer engineering, 43(03), 747-761.
[9] Jin, X. B., Zheng, W. Z., Kong, J. L., Wang, X. Y., & Lin, S. (2021). Deep-learning forecasting method for electric power load via attention-based encoder-decoder with bayesian optimization. Energies, 14(6), 1596.
[10] Zhuo, C., Long-Xiang, S., & University, Z. (2018). Short-term electrical load forecasting based on deep learning lstm networks. Electronic Technology, 13(11), 249-267.
[11] Yan, X., & Zhu, H. (2023). A kernel-free fuzzy support vector machine with universum. Journal of Industrial and Management Optimization, 19(1), 282-299.
[12] Pengnian Qi,Yulun Liao,Biao Qin. (2023). Survey on deep learning for chinese named entity recognition. Journal of Chinese Computer Systems, 44(9), 1857-1868.
[13] Liu, P., Ahmad, S., Abdullah, S., & Mohammed M. Al-Shomrani. (2022). A new approach to three-way decisions making based on fractional fuzzy decision-theoretical rough set. International Journal of Intelligent Systems, 37(3), 2428-2457.
[14] Tie, J., Lei, X., & Pan, Y. (2022). Metabolite-disease association prediction algorithm combining deepwalk and random forest. Tsinghua Science and Technology, 27(1), 58-67.
[15] Yang, L. I., Wang, Q. Y., Tian, Q. H., Qi, A. N., Yang, Y. T., & Zhang, J. C., et al. (2024). Prediction of renal function by urinary lead and cadmium — based on classification decision tree and logistic regression model*. Biomedical and Environmental Sciences, 37(3), 331-335.
[16] Lyu, J., Bi, D. J., Liu, B., Yi, G., Zheng, X. P., & Li, X. F., et al. (2023). Compressive near-field millimeter wave imaging algorithm based on gini index and total variation mixed regularization, 21(1), 65-74.
[17] Tie, J., Lei, X., & Pan, Y. (2022). Metabolite-disease association prediction algorithm combining deepwalk and random forest. Tsinghua Science and Technology, 27(1), 58-67.
[18] Du, C., Du, C., Huang, L., Wang, H., & He, H. (2022). Structured neural decoding with multitask transfer learning of deep neural network representations. IEEE transactions on neural networks and learning systems, 33(2), 600-614.
[19] Kong, Q., & Chang, X. (2022). Rough set model based on variable universe, 7(3), 9.
[20] Al-Tameemi, I. K. S., Feizi-Derakhshi, M. R., Pashazadeh, S., & Asadpour, M. (2023). Multi-model fusion framework using deep learning for visual-textual sentiment classification, 76(8), 2145-2177.

Edited by: Hailong Li
Special issue on: Deep Learning in Healthcare
Received: May 11, 2024
Accepted: Jun 21, 2024