2025 5th International Conference on Artificial Intelligence and Industrial Technology Applications (AIITA)

Research on CNN-LSTM based reprocessing process state prediction model

1st Limei Zhao*
China Nuclear Power Engineering Co.LTD
Beijing, China
zlm206926@163.com

2nd Changfu Ding
China Nuclear Power Engineering Co.LTD
Beijing, China
dingcf@cnpe.cc

3rd Bo Zhang
China Nuclear Power Engineering Co.LTD
Beijing, China
Xi'an Jiaotong University
Xi'an, China
zhangbob@cnpe.cc

4th Shihai Ma
China Nuclear Power Engineering Co.LTD
Beijing, China
Tianjin University
Tianjin, China
mash@cnpe.cc

***Abstract*— Aiming at the situation that the control system does not have a perfect interlocking protection mechanism when abnormal working conditions occur in the reprocessing process, or the operator cannot make correct judgement in time, which leads to the occurrence of accidents, a CNN-LSTM-based reprocessing process state prediction model is proposed. The wet purification system process data in the reprocessing process is taken as the research object. Firstly, the data preprocessing is carried out to address the problem of missing data, and then the structural design and training of the CNN-LSTM prediction model is completed, and finally, a part of the data is selected to test the trained model, and the consistency of the prediction results with the actual results is evaluated. The results show that the prediction results have a high degree of fit to the actual results. Therefore, the CNN-LSTM-based reprocessing process state prediction model proposed in this paper can accurately predict the state of the system.**

Keywords—CNN-LSTM, prediction mode, wet purification system

I. INTRODUCTION

Nuclear chemical process conditions often involve strong radioactivity, flammable and explosive, high temperature and pressure, toxic and harmful, etc. The detection of process parameters is of great significance for the identification of abnormalities and ensuring the smooth operation of the device. In actual engineering applications, due to the complexity of the process, the process parameters are coupled with each other, usually set a large number of alarm thresholds and alarm values to set a large margin to meet the safety requirements of the operation process, but this situation will have a certain impact on the process productivity. However, this situation will have a certain impact on the production efficiency of the process. Moreover, once a failure occurs, a large number of alarm messages will appear in a short period of time for the associated operating parameters, and the operator will not be able to judge the cause of the alarm in a short period of time, which will affect the production or lead to operational errors. If we can predict the trend of parameter changes before the occurrence of abnormal conditions, identify abnormal conditions in advance, and achieve auxiliary judgement and optimisation of DCS interlocking mechanism, we can not only effectively prevent accidents from occurring, but also give operators enough time to investigate the cause of the failure. Therefore, the process parameter prediction model is of great significance in monitoring the status of operating devices, identifying abnormal conditions, ensuring system installation and guiding operators.

In recent years, with the continuous development of artificial intelligence algorithms, neural network for state prediction has become a proven method. With the gradual deepening of neural network research, Shen[1] et al.proposed an average deviation control (ADC) state prediction control algorithm based on the pre-conducting steam temperature, and added the ADC controller into the conventional PID main steam temperature control system;A new prognostic method is developed in this paper using adaptive neuro-fuzzy inference systems (ANFISs) and high-order particle filtering[2];Yu[3] et al.proposed a medium and long-term load variable weight combination prediction model based on grey theory, through the analysis of the direction of the load curve, the combination of different trends in the stage of the combination of different grey models were established for prediction, through the cumulative residuals based on the variable weight combination prediction to get the final results;Ye[4] et al.proposes a combined EMD-SVM short-term wind power prediction method based on empirical mode decomposition (EMD) and support vector machine (SVM), which can track the changes of wind power well and improve the accuracy of short-term wind power prediction effectively. Wang[5] et al.proposed to use adaptive BP neural network to predict the parameters of experimental fast reactor, which verified the feasibility and advantages of adaptive BP neural network predictive parameter research method in the field of fast reactor; Zhang[6] et al. proposed to use multi-feature fusion multi-gait prediction model with long and short-term memory to model and predict the data of steam pressure sensor and the test results verified the applicability of the long-short-term memory network (LSTM) model in the prediction of the operating state of a nuclear power plant; Jiang[7] et al. proposed a -support vector regression machine (-SVR)-based prediction of core power under accidental conditions; Ji[8] et al. used a coupled optimisation algorithm of long-short-term memory neural networks to improve the prediction of abnormal operating conditions.Li[9] et al.uses CNN-LSTM technique to analyse the whole flow field in an air cyclone centrifugal classifier and predicts pressure. Compared with other methods mentioned above,LSTM in deep learning is more advantageous in solving the problem of long sequence data and its computational power is better than other schemes[10-12]. In this paper, based on the CNN- LSTM deep learning model, a high-precision and fast prediction analysis of the outlet flue gas temperature of the emergency cooler unit in the wet purification system is carried out, and compared with the actual operating parameters to verify the accuracy of the analysis model.

II. MODEL STRUCTURE

Convolutional neural network (CNN) is a deep learning algorithm with a convolutional structure that includes a convolutional layer, a pooling layer, and a fully connected layer[13-15], the convolutional layer is used to extract local features of the input parameters, the pooling layer can reduce the amount of data processed while retaining the useful information and is used to reduce the dimensionality in order to reduce the complexity of the data, and the pooling layer is used for the final classification or regression.CNNs have strong ability of local feature extraction through the convolution layer, and reduce the dimension and size of the data through the pooling layer. CNN extracts data features through convolutional layer, reduces data dimension and size through pooling layer, and has powerful local feature extraction capability.In CNN, the mathematical model of the convolutional layer can be represented as:

$$y_{ij} = \sum_{k=1}^{K} x_{ik} * w_{kj} + b_j \tag{1}$$

Where, x_{ik} denotes the feature value at the i th position of the input image, w_{kj} denotes the weight of the convolution kernel, b_j denotes the bias term, and y_{ij} denotes the feature value after convolution. The mathematical model of the pooling layer can be expressed as:

$$y_{ij} = \max_k (x_{ik}) \tag{2}$$

Where x_{ik} denotes the feature value at the i th position of the input image and y_{ij} denotes the pooled feature value. Long Short-Term Memory Network (LSTM) adds cell states and gate structure to the traditional Recursive Deep Network (RNN), where the gate structure includes three gating activation functions, namely, input gate, forgetting gate, and output gate[16-17].RNN is characterised by long-distance dependency, and suffers from the problem of gradient vanishing.LSTM, as a deep learning algorithm derived from RNN, is able to efficiently capture long-term dependency, and shows good capability in analysing and prediction of time series data shows better ability. The mathematical model of LSTM can be expressed as:

$$i_t = \sigma(W_{xi} x_t + W_{hi} h_{t-1} + b_i) \tag{3}$$

$$f_t = \sigma(W_{xf} x_t + W_{hf} h_{t-1} + b_f) \tag{4}$$

$$g_t = \tanh(W_{xg} x_t + W_{hg} h_{t-1} + b_g) \tag{5}$$

$$o_t = \sigma(W_{xo} x_t + W_{ho} h_{t-1} + b_o) \tag{6}$$

$$c_t = f_t * c_{t-1} + i_t * g_t \tag{7}$$

$$h_t = o_t * tanh(c_t) \tag{8}$$

Where x_t denotes the eigenvalue at the t th time step of the input sequence, h_t denotes the hidden state of the LSTM, and c_t denotes the cellular state of the LSTM. σ denotes the sigmoid function, and $tanh$ denotes the hyperbolic tangent function. W_{xi} , W_{hi} , W_{xf} , W_{hf} , W_{xg} , W_{hg} , W_{xo} , W_{ho} denote the weight matrix, b_i , b_f , b_g , b_o denotes the bias vector. Its network structure is shown in Figure1.

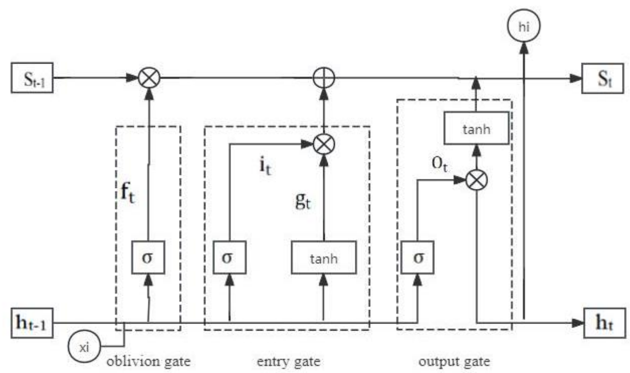

Figure1: LSTM Network Structure

CNN-LSTM is a deep learning algorithm that combines CNN and LSTM[18-20]. This combination can effectively capture both local features and long-term dependencies, which allows for more effective extraction of state features and prediction from complex time-series data.

III. MODEL TRAINING

In this paper, the process operation data of the wet purification system in the post-treatment process is taken as the research object, and the CNN-LSTM model is trained in order to achieve the prediction of the flue gas temperature in the wet purification system two minutes afterward after the cooling of the rapid cooler and scrubber tower by the actual generated time series data. Six parameters related to the tower outlet temperature are selected as input parameters: the inlet temperature of the emergency cooler, the scrubber tower pressure, the scrubber liquid flow rate of the emergency cooler, the scrubber liquid flow rate of the scrubber tower, the scrubber liquid outlet flow rate of the scrubber pump, and the scrubber tower outlet temperature, while the scrubber tower outlet temperature is also a predicted parameter. The training process of the model is shown in Figure 2 below.

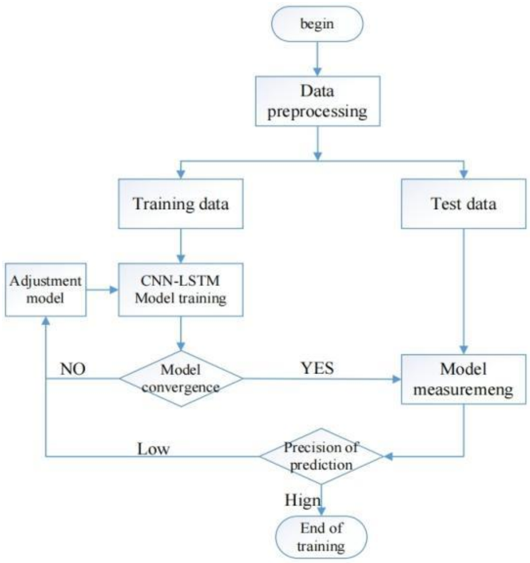

Figure 2: Model Training Process

In order to speedup the model training and improve the accuracy of the model, it is first necessary to pre-process the operational data of the system, including operations such as noise removal, standardisation and normalisation, to ensure the quality and consistency of the input data.

The CNN-LSTM algorithm model of this experiment consists of 1 layer CNN and 1 layer LSTM network, CNN is used to extract spatial features in time series data and LSTM network is used to capture temporal features in time series data. The specific experimental parameter settings are shown in TABLE I.

TABLE I. PARAMETER SETTINGS FOR CNN-LSTM

Parameter name	CNN-LSTM model
batch_size	32
Learning_rate	0.01
Num_epochs	100
Input_shape	(5861,10,6)
Num_calss	(5861,)
Learning device	Adam
Preprocess	Normalisation
Input layer	Input Functions
Coding layer	Conv 1D, MaxPooling1D Functions
Predictive layer	The scrubber tower outlet temperature
Ponding layer	1

Since the processing is one-dimensional sequence data, the CNN neural network uses a one-dimensional convolu- tional layer (Conv1d) to extract features from the original data, the size of the Conv1d convolutional kernel is 1×6, 6 is the dimension of the input original data, the number of convolutional kernels is 16, and the moving step is 1. The output of the convolutional layer is transformed nonlinearly by the activation function sigmoid. The features are extracted and then fused into 1×16 fusion features by convolution. The pooling layer uses a one-dimensional maximum pooling operation, MaxPool1d, for processing time series data.The main parameters of MaxPool1d include kernel_size (the size of the sliding window), stride (the step size of the sliding window), padding (negative infinity padding implicitly added to the sides), dilation (the (step length between elements in the sliding window). In this paper, we design the values of each parameter as MaxPool1d(kernel_size=5, stride=5, padding=0, dilation=1). During the training process, in order to prevent overfitting and enhance the generalisation ability of the model, the regularisation method is used to randomly set certain elements of the input tensor to zero with a probability of 0.01, and these operations do not directly modify the original data, but return a new tensor. In this paper, a bidirectional LSTM isused to receive the output data from the CNN, the number of nerve cells (cells) in each layer is 16, and the activation function is the Tanh function, and each nerve cell outputs a hidden state volume after processing, and the output after processing by the LSTM neural network is a 1×16 vector. After that the activation function layer tanh function is used and the output of the LSTM layer is nonlinearly transformed. Then a linear layer with 32 input features and 1 output feature is defined by a fully connected layer which will contain a learnable bias term and the output of the fully connected layer is transformed by another tanh transform using the last activation function layer. The final value after back-normalisation is the predicted value.The mean square error (MSE) is constructed as the model loss function, MSE is widely used in deep learning for regression tasks tooptimise the model parameters by minimising the difference between the predicted and true values. The Adam optimisation algorithm is used to find the optimal value of the loss function, and each layer of weights and bias values are optimised through layer-by-layer training until the loss function converges to a stable value.

IV. MODEL PREDICTION RESULTS AND ANALYSIS

The experiment aims at predicting the flue gas temperature at the outlet of the rapid cooler and scrubber tower in the wet purification system in the reprocessing process using a CNN-LSTM deep learning model.The pytorch deep learning framework was used and the model was trained using Adam optimiser and MSE loss function.Its network structure is shown in Figure 3.

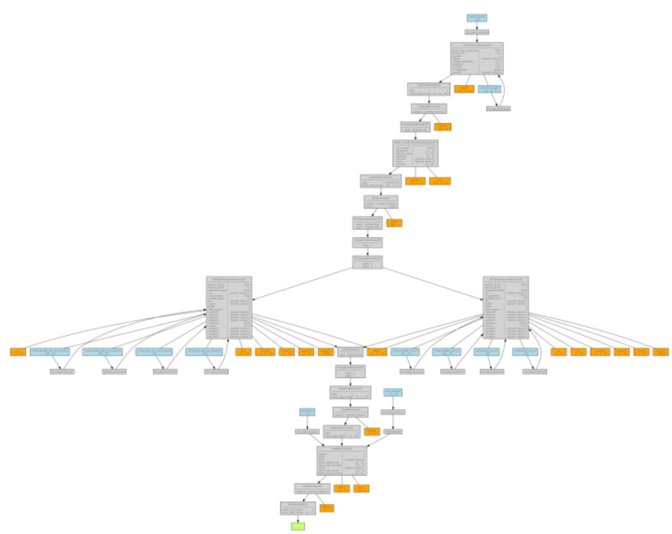

Figure 3: CNN-LSTM Network Structure

Select 90% of the data as the training set, train the model until the loss function converges, and then use the other 10% of the data as the test set to get the predicted values and compare with the actual values, the results are shown in Figure 4. The red curve in the figure is the predicted temperature curve, which is the exit flue gas temperature in the last two minutes predicted from the historical data of the first ten minutes. The blue curve is the actual temperature curve. The experimental results show that CNN-LSTM can predict the exit flue gas temperature data of the following two minutes by any ten-minute historical data, and the correlation between the predicted value and the measured value is high (R2>0.85). The root mean square value (RMSE), absolute mean (MAE), and mean absolute percentage error (MAPE) evaluation indexes were used to evaluate the accuracy of the model prediction results. The values of each evaluation index are shown in TABLE II below.

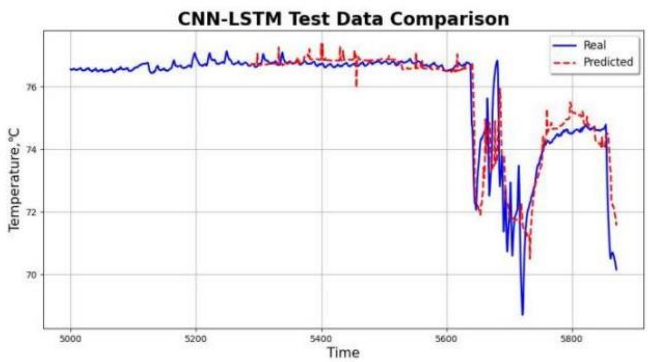

Figure 4: Comparison Results between Real Values and Predicted Values under CNN-LSTM Model

TABLE II. EVALUATION INDICATORS

Evaluation indicators			
R^2	RMSE	MAE	MAPE
0.8628	0.6859	0.3846	0.5207%

V. CONCLUSION

In this paper, CNN-LSTM model is used to predict the exit flue gas temperature in the wet cleaning system in the reprocessing process. The main function of this model is to extract features from sequence data and perform classification. Firstly, the initial feature extraction and nonlinear transformation are performed by a 1D convolutional layer and an activation function; then, the overfitting is reduced by a maximum pooling and dropout layer; then, the temporal information in the sequences is further captured by a bi-directional LSTM; and finally, the final classification task is accomplished by a fully-connected layer and an activation function. After the model training is completed, the real-time running data can be fed into the established fault diagnosis model for prediction. The results show that the model has high prediction accuracy and efficiency, and can accurately predict the exit flue gas temperature of the wet evolution system, which in turn can prove that this model, after training, can also be applied to other nuclear chemical processes that have a need for parameter prediction.

Therefore, in practical application, the model can predict the trend of detection parameters over a period of time through historical data, which can detect potential failures and take preventive measures in time, thus reducing downtime and maintenance costs. In the subsequent application of the model, in order to achieve better prediction results, the structure and parameters of the model will be further optimised, and efforts will be made to improve and adjust the model, for example, to further study the effect of prediction time on the prediction accuracy of the system as well as the processing efficiency, and also through the introduction of methods such as the attention mechanism, in order to enhance its prediction performance and obtain more accurate prediction results.

REFERENCES

[1] SHEN Jialin,HE Jianjia. A main steam temperature control of thermal power plant with ADC state predictive control algorithm[J]. Thermal Power Generation, 2010, 39(12):4. DOI:10.3969/j.issn. 1002-3364.2010. 12.040.

[2] Chen C, Zhang B, Vachtsevanos G ,et al.Machine Condition Prediction Based on Adaptive Neuro–Fuzzy and High-Order Particle Filter-ing[J].IEEE Transactions on Industrial Electronics,2011,58(9):4353-4364.DOI:10. 1109/TIE.2010.2098369.

[3] YU Jianming,YAN Fei,YANG Wenyu,et al. Variable-weight grey combinatorial forecasting model for medium-and long-term electricity load[J]. Grid Technology, 2005, 29(17):4. DOI:10.3321/j.issn:1000-3673.2005. 17.006.

[4] Ye Lin,Liu Peng. Short-term wind power mix prediction model based on empirical modal decomposition and support vector machine[J]. Chinese Journal of Electrical Engineering, 2011, 31(31):7. DOI:CNKI:SUN:ZGDC.0.2011-31-013.

[5] WANG Dongdong, YANG Hongyi, WANG Duan, et al. Research on adaptive BP neural network prediction method for thermal parameters of Chinese experimental fast reactor[J]. Atomic Energy Science and Technology, 2020, 54(10): 1809- 1816.

[6] Zhang Si-yuan, Lu Xiongyu, Zeng Hui, et al. Multi-feature fusion multi-step state prediction for nuclear power sensors based on LSTM[J]. Nuclear Power Engineering, 2021,42(4): 208-213.

[7] JIANG B T, HUANG X B, WESLEY H J, et al. Reactor power prediction under accidental operating conditions based on U-support vector machines[J]. Nuclear Power Engineering, 2019, 40(6): 105- 108.

[8] Ji N, Yang JK, Zhao P, Wang K et al. Prediction of accident parameters in lead-bismuth reactors coupled with multivariate LSTM and optimisation algorithm[J]. Nuclear Power Engineering, 2023, 44(5): 64-70.

[9] Li W, Li X, Yuan J.Pressure prediction for air cyclone centrifugal classifier based on CNN-LSTM enhanced by attention mechanism[J].Chemical Engineering Research&Design: Transactions of the Institution of Chemical Engineers, 2024:205.DOI:10. 1016/j.cherd.2024.04.045.

[10] Heng F, Gao J, Xu R,et al.Multiaxial fatigue life prediction for various metallic materials based on the hybrid CNN-LSTM neural network[J].Fatigue & Fracture of Engineering Materials & Structures, 2023.DOI:10. 1111/ffe. 13977.

[11] Sun Y , Zhou Q , Sun L ,et al.CNN–LSTM–AM: A power prediction model for offshore wind turbines[J].Ocean Engineering, 2024, 301.DOI:10. 1016/j.oceaneng.2024. 117598.

[12] Ma Liangyu,Lu Ruomeng. Optimisation of CNN-LSTM-SEnet prediction model with SSA for wind turbine fault warning[J]. Power Science and Engineering, 2024(6).

[13] Shin H C , Roth H R , Gao M ,et al.Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning[J].IEEE Transactions on Medical Imaging, 2016, 35(5):1285- 1298.DOI:10. 1109/TMI.2016.2528162.

[14] Radford A , Metz L , Chintala S .Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Net- works[J].Computer ence, 2015.DOI:10.48550/arXiv. 1511.06434.

[15] ZHOU Chaomian,LIU Mingping,WANG Jingwei. Research on water quality prediction model based on CNN-LSTM[J]. Hydropower Energy Science, 2021, 39(3):4.

[16] Gers F A , Eck D ,Jürgen Schmidhuber.Applying LSTM to Time Series Predictable through Time-Window Approaches[C]//International

Conference on Artificial Neural Networks.Springer, Berlin, Heidelberg, 2001.DOI:10. 1007/3-540-44668-0_93.

[17] Gers, Felix A ,Schmidhuber,et al.Learning to Forget: Continual Prediction with LSTM.[J].Neural Computation, 2000.DOI:10. 1162/089976600300015015.

[18] QIN Yu,OUYANG Changyue,FANG Peng. Prediction of reservoir carbon dioxide flux based on CNN-LSTM model and small sample data[J]. Journal of Chongqing Jiaotong University: Natural Science Edition, 2022(006):041.

[19] KONG, J., LI, YABIN, ZHANG, SHIYU, et al. Research on trajectory prediction based on CNN-LSTM-attention model[J]. Aviation Computing Technology, 2023, 53(1):1-5.

[20] CHEN Yibin,LU Jiaxuan,ZHANG Shuming,et al. A deep learning model for PM_(2.5) concentration prediction based on hybrid CNN-LSTM structure[J]. Information and Computer, 2022(004):034.