Transactions on Engineering and Computing Sciences - Vol. 11, No. 5
Publication Date: October 25, 2023
DOI:10.14738/tecs.115.15411.
Yidan, X., Shaolin, H., & Guotao, Y. (2023). Analysis and Improvement Approach of the Impact of Data Disturbance on LSTM Prediction Algorithm. Transactions on Engineering and Computing Sciences, 11(5). 01-15.



Analysis and Improvement Approach of the Impact of Data Disturbance on LSTM Prediction Algorithm

Xin Yidan
School of Automation and Electrical Engineering,
Xi'an University of Technology, Xi'an, China

Hu Shaolin
Automation School, Guangdong University of
Petrochemical Technology, Maoming, China

Yang Guotao
School of Automation and Electrical Engineering,
Xi'an University of Technology, Xi'an, China

ABSTRACT

If the sampling data with noise or outliers are used to train the long-short-term memory (short as LSTM) network, whether the perturbations and outliers in sampling data affect the training performance and prediction accuracy of LSTM networks is a key problem. This paper analyzed the impact on the LSTM prediction effect when using perturbations and isolated/patchy outliers involved data for network training and prediction. The results showed that the prediction accuracy decreases as the amplitude of the perturbations and the range of outliers increase. In order to overcome this effect, an improved method of Pre-set Outliers Tolerant Filter is proposed, and an Outliers-Tolerant Multi-LSTM model, in short, the OTML model, is established. The prediction effect of the proposed model is compared with that of the LSTM model without the filter and that of the LSTM model with a mean filter. Comparison results showed that the OTML prediction model proposed in this paper can eliminate the influence of random noise and isolated outliers thoroughly. When meeting the patchy outliers whose width is smaller than the radius of the window, the OTML prediction model can filter them out too, so as to realize the high-fidelity prediction of the data.

Keywords: LSTM, Outlier, Fault-tolerant filtering algorithm, Fault tolerance.

INTRODUCTION

In practical application scenarios, strong interference factors in the external environment during the sampling task usually lead to randomly distributed noise or outliers in the sampled data [1]-[3], and when analyzing the working conditions and predicting the future trend based on the sampled data, noise and outliers reduction will be carried out to promote the accuracy and credibility of the prediction results [4]. Conventional methods are mean filtering, least squares filtering, linear interpolation filtering, Kalman filtering, etc [5]. Among the above algorithms, the mean filtering algorithm may lead to large-scale fluctuations when facing

continuous signals and may lead to a poor smoothing effect for signals with high peaks [6]. The least squares filtering algorithm spends too much time and requires many iterations during the implementation of the algorithm, thus resulting in a high degree of computational complexity [7]. Linear interpolation filtering is less effective for cases with a large amount of missing data, which may lead to a large estimation error. The Kalman filtering requires good assumptions about the state equation and the observation equation, which is more complex, and the computational cost is higher [8]. Therefore, how to completely and efficiently eliminate the influence on LSTM prediction, which is caused by the random noise and outliers contained in the data, is very important [9]. In this paper, an improved method based on outlier-tolerant filtering is established, which realized the prediction of real trends in data by introducing an outlier-tolerant filtering layer into the multilayer LSTM network to form the Outliers-Tolerant Multi-LSTM model, in short, the OTML model.

The specific structure of the paper is as follows: In section 2, the impact of data quality on LSTM prediction accuracy is analyzed using data containing random noise and data containing isolated/patchy outliers. In section 3, an improved method based on an outlier-tolerant filtering algorithm is established. In section 4, experimental data are used to validate the method proposed in this paper, and comparative experiments are set up. In section 5, the utility and benefits of the OTML predictive model are summarized.

IMPACT OF DATA IMPURITIES ON LSTM PREDICTION

Theoretical Analysis

LSTM network model is a kind of gated deep recurrent neural network, which can realize the long-term memory of the time series, and can predict the time series data of the next moment according to the current moment and the past moment data comprehensively, also can fully learn the regularity of the data, which has a good prediction effect for the time series data [10]-[13].

The hidden layer of the LSTM network consists of four interacting neural network layers: the input gate, the forgetting gate, the internal state, and the output layer [14]. Using $\tilde{y}(t_i)$ as input, the forward learning process of LSTM is as follows.

$$i_{t_i} = sigmoid(W_{yi}\tilde{y}(t_i) + W_{hi}h_{t-1} + b_i) \quad (1)$$

$$f_{t_i} = sigmoid(W_{yf}\tilde{y}(t_i) + W_{hf}h_{t-1} + b_f) \quad (2)$$

$$o_{t_i} = sigmoid(W_{yo}\tilde{y}(t_i) + W_{ho}h_{t-1} + b_o) \quad (3)$$

$$c_{t_i} = f_{t_i t} \odot c_{t_{i-1}} + i_{t_i} \odot tanh(W_{yc}\tilde{y}(t_i) + W_{hc}h_{t-1} + b_c) \quad (4)$$

$$h_{t_i} = o_{t_i} \odot tanh(c_{t_i}) \quad (5)$$

Where i_{t_i} denotes the input gate, f_{t_i} denotes the forgetting state, c_{t_i} denotes the internal state, o_{t_i} denotes the output gate, b is the bias and W is the weight parameters, h_{t_i} is the final output.

The bias and weight parameters in the network model are updated using the Back Propagation Through Time (BPTT) algorithm. The LSTM is trained by time series $\tilde{y}t_1), \tilde{y}(t_2), \ldots, \tilde{y}(t_n)$, n is the length of the time series. The data is divided into training and test sets at the ratio of 7:3 in the input layer, and the training set is normalized using Z-score normalization to obtain the normalized training set $\bar{y}(t_0), \bar{y}(t_1), \ldots, \bar{y}(t_n)$. Before inputting the training set into the network, the data in the training set should be transformed into a three-dimensional format, and the transformed training set data should be partitioned into data blocks through a sliding window for training the network, and the mean square error should be chosen to compute the error loss between the model output value and the theoretical output value, and the point-by-point iterative computation method is used in the prediction process of the LSTM network, notate the training-completed network as trained model. Input the $\bar{y}(t_0), \bar{y}(t_1), \ldots, \bar{y}(t_n)$ into the trained model to get the predicted series $\hat{y}(t_0), \hat{y}(t_1), \ldots, \hat{y}(t_n)$.

Due to the complexity of the operating state and measurement environment in which the data acquisition target is located, the data usually used to train the LSTM network will contain some impurities, including large random perturbations and a small portion of anomalous data that seriously deviates from the main body of the data, which can be expressed as follows:

$$y(t) = s(t) + \varepsilon(t) + o(t) \tag{6}$$

In the above equation, $s(t)$ represents the sequence of true values, $\varepsilon(t)$ is the random perturbation, and $o(t)$ is the anomalous data sequences. It can be seen from the equation (6) that both $\varepsilon(t)$ and $o(t)$ may cause the LSTM focus on the extreme values in the dataset during training excessively, which may lead to the overfitting and the generalization ability to new data deteriorates. It is not enough to analyze the possible effects only theoretically, specifically, this paper intends to separately analyze the effects of random perturbations with different variance and different amounts of outliers on the prediction accuracy of LSTM.

Effects of Perturbations with Different Variances on LSTM Prediction

The structure of LSTM networks is very complex; thus, it is difficult to theoretically analyze the effect of random perturbation on LSTM. Therefore, this section adopts the method of experimental analysis, the measured data of heating chamber temperature provided by a petrochemical base is used as the original data, and random perturbations with different variances are added to the original data to generate experimental simulation data. Prediction experiments are conducted based on this simulation data, and then analyze the effects of different degrees of random perturbations on the prediction accuracy of LSTM.

Specifically, the experimental simulation data are as follows:(1) A random noise with a variance size of 1 is added to the original data $y_1(t_i)$ to obtain the experimental data $y_2(t_i)$.(2) A random noise with a variance size of 9 is added to the original data $y_1(t_i)$ to obtain the experimental data $y_3(t_i)$. $y_1(t_i)$, $y_2(t_i)$ and $y_3(t_i)$ are shown in subfigures (a.1) to (a.3) in Fig. (1).

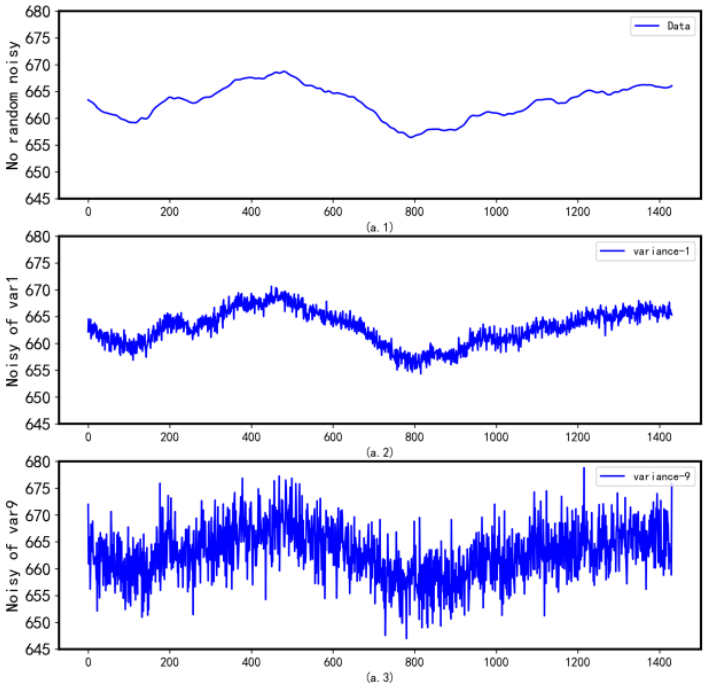


Fig. 1: Experimental data containing random noise of different variances

Separately, $y_1(t_i)$, $y_2(t_i)$ and $y_3(t_i)$ are divided into training and test sets according to the ratio of 1200:241, and 30% of the data in the training set is used as the validation set, based on which three different LSTM prediction models are trained respectively, and these three models are used to predict the original data $y_1(t_i)$. The prediction results are noted as $\hat{y}_1(t_i)$, $\hat{y}_2(t_i)$ and $\hat{y}_3(t_i)$, and the prediction residuals between each prediction result and the original data are noted as $\Delta y_1(t_i)$, $\Delta y_2(t_i)$ and $\Delta y_3(t_i)$.

Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) are obtained based on the residuals of each prediction, and the above metrics are used to evaluate the prediction accuracy of LSTM, when the prediction accuracy is higher, the value of each metric is smaller, on the contrary, the prediction accuracy is lower, the metrics are bigger.

Through iterations, each LSTM prediction model is tuned to the optimal, and each prediction result and residual sequence are obtained based on the optimal model, and the experimental results are shown in Figure 2 below.

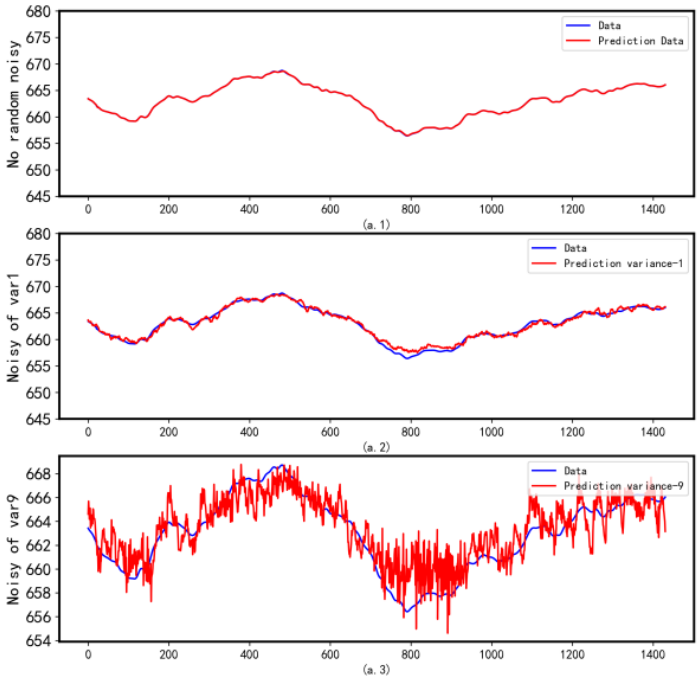


Fig.2: (a) prediction result

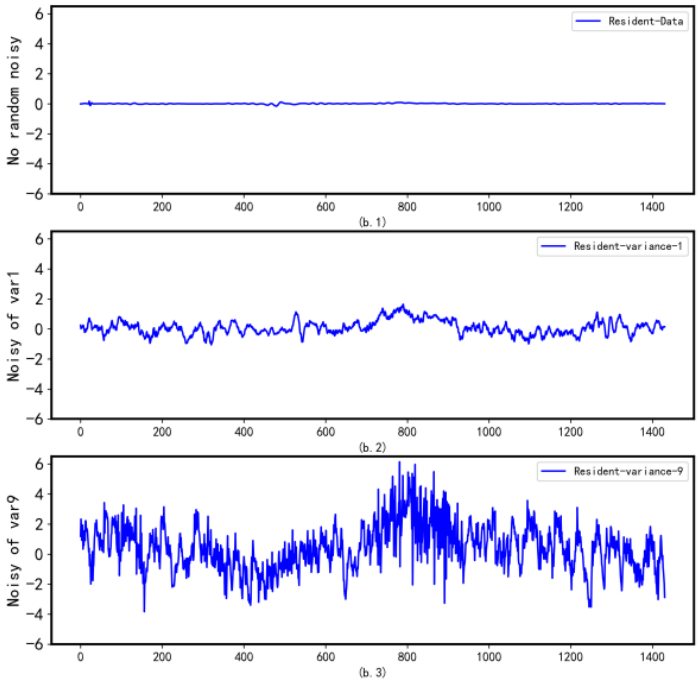


Fig.2: (b) prediction resident

Fig. 2: LSTM prediction results on data with different variance

In the above figure, subfigure (a.1) shows the prediction curve of $y_1(t_i)$ after training the network based on $y_1(t_i)$. Subfigure (a.2) shows the prediction curve of $y_1(t_i)$ after training the network based on $y_2(t_i)$. Subfigure (a.3) shows the prediction curve of $y_1(t_i)$ after training the network based on $y_3(t_i)$. And subfigures (b.1) to (b.3) showed the corresponding prediction

residents. As can be seen from each figure, when the network is trained and predicted based on $y_1(t_i)$, the predicted values roughly coincide exactly with the true values, and as the variance of random perturbations arise, the difference between the predicted results and the true values gradually increases. And the range of the predicted residuals varies from near 0 to within the interval [-5,5].

Table 1 gives the prediction performance metrics of LSTM on each data using RMSE, MAE, and MAPE as metrics for evaluating prediction accuracy.

Table 1: Predictive performance metrics of different data trained LSTM

Training Data	RMSE	MAPE	MAE
$y_1(t_i)$	0.00121	0.00532	0.02136
$y_2(t_i)$	0.08647	0.0054	0.23526
$y_3(t_i)$	0.14979	0.0054	0.29819

As can be seen in Table 1, the smallest values of the indicators are obtained when training and predicting the network based on the original data $y_1(t_i)$, and the worst prediction performance indicators are obtained when training the network based on $y_3(t_i)$.

From above Figure 1 and Table 1, it can be seen that random perturbations have a certain impact on the prediction of LSTM. Specifically, the prediction residuals shown in Figure 1(b) showed that as the variance of random noise increases, the prediction residuals gradually increase. Also, the evaluation metrics shown in Table 1 indicate that the network trained using data containing random noise with a variance of 9 has a larger variance than when the random noise which has a variance of 1, i.e., the prediction accuracy deteriorates. Therefore, it can be concluded that the existence of random noise will not only make the residuals of LSTM prediction increase, and the increase in variance will exacerbate the degree of deformation of the prediction results.

Effect of Isolated/Patchy Outliers on LSTM Prediction

Due to the complexity of the complexity of the structure, it is difficult to carry out the theoretical derivation of LSTM. This section adopts the method of data experimental analysis, using the heating furnace temperature sensing data as the original data $y_1(t_i)$, the isolated outliers and patchy outliers are respectively added to the original data to generate the experimental simulation data. Prediction experiments are conducted based on these simulation data, then the effect of isolated/patchy outliers on the prediction accuracy of LSTM is analyzed.

Specifically, the experimental simulation data are as follows:(1)6 isolated outliers are introduced into $y_1(t_i)$ to obtain the experimental data $y^*{}_2(t_i)$. (2) 5 groups of patchy outliers with a width of 6 are introduced into $y_1(t_i)$ to obtain the experimental data $y^*{}_3(t_i)$. $y_1(t_i)$, $y^*{}_2(t_i)$ and $y^*{}_3(t_i)$ are shown in subfigures (a.1) to (a.3) in Fig. (3).

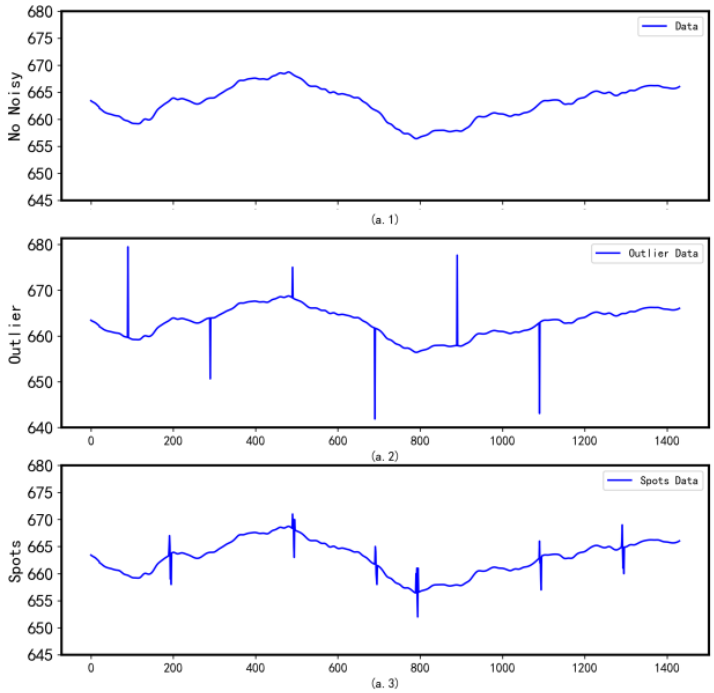


Fig. 3: Experimental data containing isolated outliers/local spots

$y_1(t_i)$, $y^*{}_2(t_i)$ and $y^*{}_3(t_i)$ are divided into training and test sets according to the ratio of 1200:241, and 30% of the data in the training set is used as the validation set, based on which three different LSTM prediction models are trained respectively, and these three models are used to predict the original data $y_1(t_i)$. $\hat{y}_1(t_i)$, $\hat{y}^*{}_2(t_i)$ and $\hat{y}^*{}_3(t_i)$ represent the prediction results, and $\Delta y_1(t_i)$、$\Delta y_2^*(t_i)$ and $\Delta y_3^*(t_i)$ represent the prediction resident.

Based on the residuals of each prediction, RMSE, MAE and MAPE are obtained, and the above indexes are used to evaluate the prediction accuracy of LSTM, when the prediction accuracy is higher, the value of each index is smaller, and vice versa, the lower the prediction accuracy is, the bigger the index is. Through many iterations, each LSTM prediction model is tuned to the optimal, and each prediction result and residual sequence are obtained based on the optimal model, and the experimental results are shown in Figure 4 below.

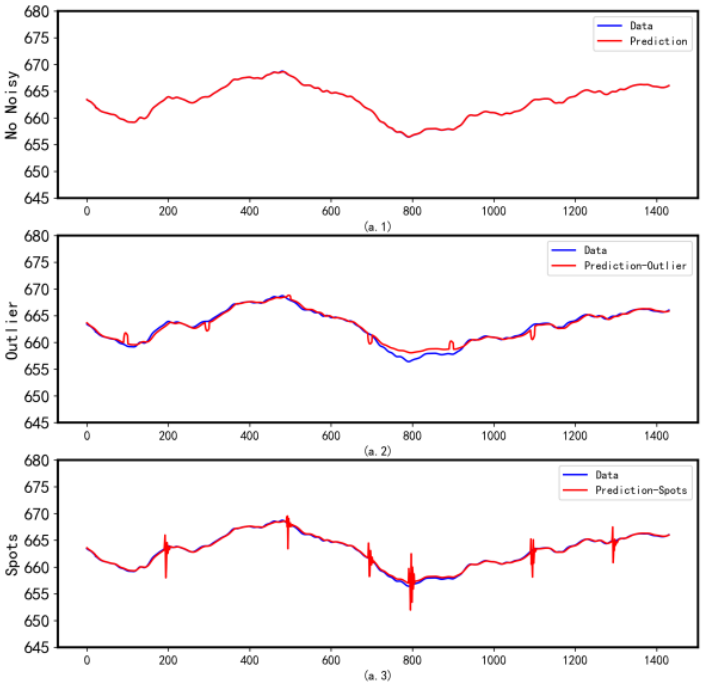


Fig.4: (a) prediction result

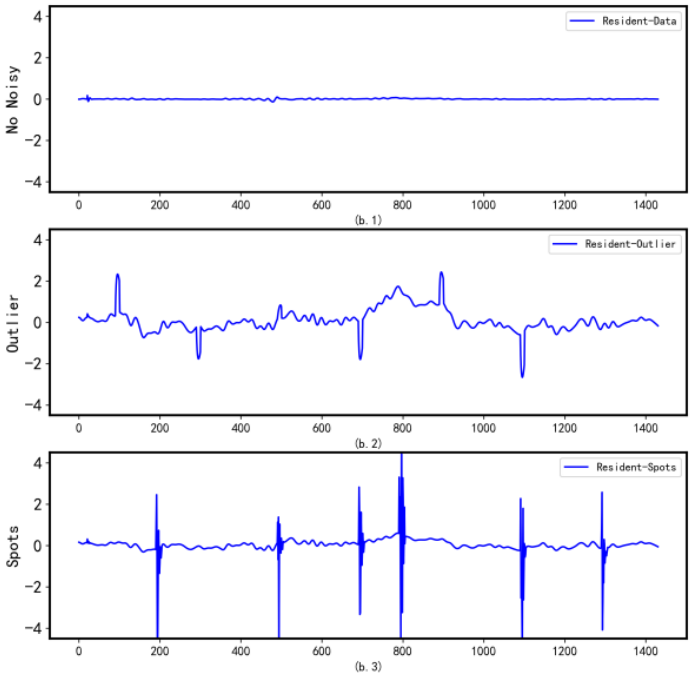


Fig.4: (b) prediction resident

Fig. 4: LSTM prediction results on data with isolated outliers/local spots

In the above figure, subfigure (a.1) shows the prediction curve of $y_1(t_i)$ after training the network based on $y_1(t_i)$. Subfigure (a.2) shows the prediction curve of $y_1(t_i)$ after training the network based on $y^*{}_2(t_i)$. Subfigure (a.3) shows the prediction curve of $y_1(t_i)$ after training the

network based on $y^*_3(t_i)$. And subfigures (b.1) to (b.3) show the corresponding prediction residents. As can be seen from the above figure, when the network is trained and predicted based on $y_1(t_i)$, the predicted value and the true value roughly coincide completely. When isolated outliers are introduced into the data, the predicted value and the original value not only have errors at the isolated outlier, but also have a tendency of deformation before the outlier appeared. When there are patchy outliers in the data, the error between the predicted value and the true value is concentrated in the place where the patchy outliers are located. The prediction error also occurs at other places, but the predicted value and the true value are basically coinciding.

Table 2 gives the prediction performance metrics of LSTM on each data using RMSE, MAE, and MAPE as metrics for evaluating prediction accuracy.

Table 2: Predictive performance metrics of different data trained LSTM

Training Data	RMSE	MAPE	MAE
$y_1(t_i)$	0.00071	0.0054	0.017041
$y^*_2(t_i)$	0.32047	0.00524	0.36263
$y^*_3(t_i)$	0.25983	0.00537	0.20209

As can be seen from Table 2, when the network is trained and predicted based on the original data, the performance indexes are the best, the prediction results of the network based on the data containing patchy outliers are slightly worse, and the prediction results of the network based on the data containing isolated outliers are the worst. From Figure 4 and Table 2, it can be seen that both isolated outliers and patchy outliers have an impact on the LSTM prediction accuracy. Specifically, when there are isolated outliers in the data, the error between the prediction results and the true value is larger. When there are local spots in the data, the error between the prediction results and the true value concentrates in the local spots.

Above all, it is clear that the impact of data quality on network prediction should not be ignored, and we must try to eliminate this effect before using the data for network training.

IMPROVED METHODS BASED ON OUTLIER-TOLERANT FILTERING

From the experiments in the previous section, it can be seen that the different magnitudes of random noise and isolated/patchy outliers contained in the data have an impact on both the training and prediction accuracy of the LSTM network. In order to eliminate this effect, an Outliers-Tolerant Multi-LSTM (OTML) based on fault-tolerant filtering is proposed in this paper, which a Q-filter based outlier-tolerant filtering algorithm is used to filter out different kinds of noises in the data, and model training and prediction are performed based on this filtered high-fidelity signal.

Design of Multilayer LSTM Prediction Model Based on Outlier-Tolerant Filtering

In order to eliminate the influence of outliers on the training effect of LSTM network as much as possible and ensure the accuracy of data prediction, this paper realizes the extraction of high-fidelity signals from the data by introducing an outlier-tolerant filtering layer before the LSTM

network. Specifically, the computational process of the error-tolerant filter layer is shown as follows.

- Initially, the raw equally spaced sampling time series is $= \{y(t_i)|t_i = t_0 + ih, i = 1, \cdots, n\}$, in which h represents the sampling interval, n represents the sampling length. Set the radius of the sliding window to k(default $k = 7$), the width of the window W is $w = 2k + 1$(default $w = 15$).
- Equivalent addition is made to both ends of the data before calculating:

$$S_f = \{y(t_i) = y(t_0), i = -1, -2, \ldots, -k\} \tag{7}$$

$$S_b = \{y(t_i) = y(t_n), i = n + 1, n + 2, \ldots, n + k\} \tag{8}$$

The extended sequence is $S_e = S_f \cup S \cup S_b$.

- Then, slide the window from left to right to sort the data in the window W from smallest to largest to get the sorted sequence $\hat{W} = \{\hat{y}(t_j), j = -k + i, \ldots, i + k\}, i = 1,2,.., n$.
- Constructing the Quartile (Quartile, Q-) mean operator:

$$Q(\{y(t_1), y(t_2), \ldots, y(t_n)\}, n) = \frac{1}{q_3 - q_1 + 1} \sum_{i=q_1}^{q_3} (y(t_i)) \tag{9}$$

In the equation, q_1 and q_3 are the upper and lower quartile numbers of W, respectively. n is the length of the sequence within the sliding window.

- Filtering S_e using the Q-mean operator to get the filtered sequence $\hat{S} = \{\hat{y}(t_i)|t_i = t_0 + ih, i = 1, \cdots, n\}$. Calculate the difference between the filtered sequence and the raw sequence to get the residual series $E = \{\tilde{y}(t_i)|\tilde{y}(t_i) = y(t_i) - \hat{y}(t_i), \ \ i = 1, \cdots, n\}$.
- Quartile smoothing filtering is performed again on the residual sequence E to get $\hat{E}$. Linearly combine the filtering results of the data sequence with the filtering results of the residual sequence to get the filtering and noise reduction results $\hat{\hat{S}} = \hat{S} + \hat{E}$.

The introduction of an outlier-tolerant filtering layer allows the LSTM network to avoid the effects of outliers in the original data, and has a good effect on the removal of spots with widths smaller than the radius of the sliding window.

Effectiveness of LSTM Prediction with Outlier-Tolerant Prefilter Layer

In order to verify the effectiveness of the mentioned method, this section uses the heating chamber temperature data provided by a petrochemical base as the original data. A random noise with a variance of 1, 4 isolated outliers and 6 patchy outliers with width of 5 are introduced into the original data to generate the experimental data, based on which the prediction effect of OTML prediction model is verified and RMSE, MAE and MAPE are adopted as the evaluation indexes of prediction effect, and the experimental data are shown in Fig. 6.

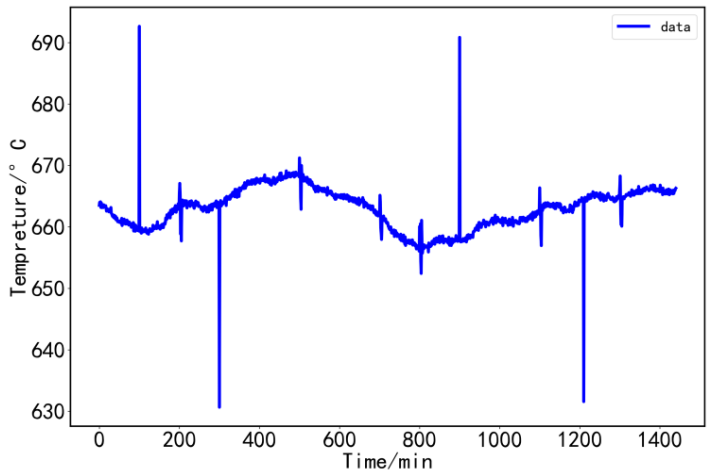


Fig.6: Experimental data containing noisy and isolated outliers/local spots

The experimental data is divided in to training set and test set at the ratio of 1200:241, and 30% of the data in the training set is used as the validation set. After training the network with the training set, the trained model is used to predict the whole segment of data, and the prediction results are shown in Fig. 7(a), the prediction residual curve is shown in Figure 7(b) below.

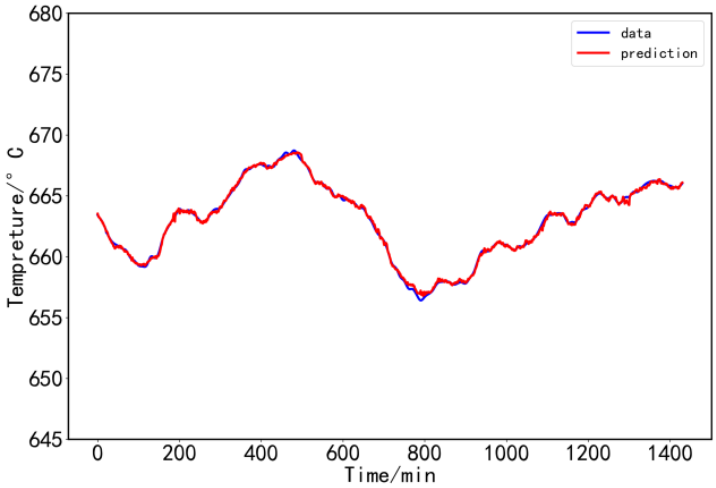


Fig.7: (a) prediction result

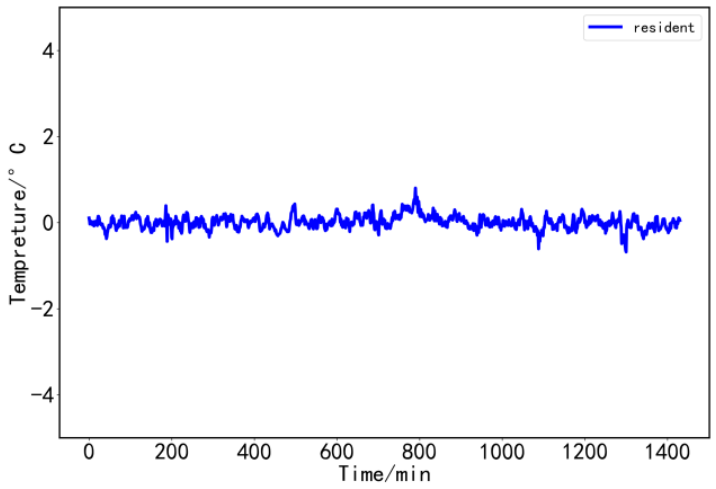


Fig.7: (b) prediction resident

Fig. 7: Prediction results of OTML on data containing noisy and isolated outliers/local spots

As can be seen from the prediction results shown in Fig. 7(a), the prediction results basically overlap with the original data when using the OTML prediction model for data that are containing noisy, isolated outliers and patchy outliers. In the prediction residuals shown in Fig. 4(b), it is clear that the prediction residuals are relatively small and are concentrated near zero. Table 3 gives the prediction performance metrics of OTML on the experimental data using RMSE, MAE, and MAPE as metrics for evaluating prediction accuracy.

Table 3: Predictive performance metrics of on different data trained LSTM

Model	RMSE	MAPE	MAE
OTML	0.01894	0.00531	0.10613

As seen in Table 3, compared to straightly predict the noisy and outliers contained data using LSTM, when the experimental data is predicted based on the OTML model, the values of each prediction index are significantly reduced.

From Fig. 7 and Table 3, it is clear that the outliers-tolerant filtering algorithm improved LSTM network can reduce the influence of random noisy, isolated/ patchy outliers on the prediction results, which verifies the effectiveness of the improved method proposed in this paper.

Comparative Experiments with Results Analysis

In order to verify that the OTML prediction model proposed in this paper has a good outlier-tolerance, this section sets up comparative experiments, and compared the prediction effect of the LSTM without noise reduction, the LSTM with mean filtering and the OTML model on the experimental data set in last section.The data set division method is consistent with the previous experiment, the prediction effect is shown in Figure 8(a) below, and the prediction residual curve is shown in Figure 8(b) below.

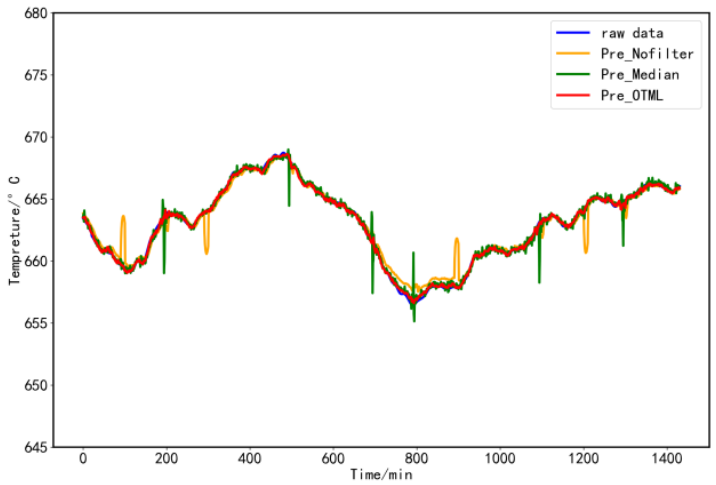


Fig.8: (a) prediction result

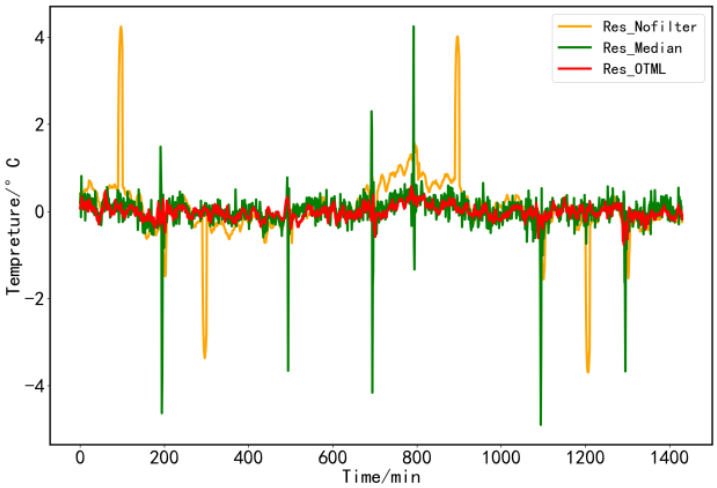


Fig.8: (b) prediction resident

Fig.8: Prediction results of different prediction model on the experimental data

In Figure 8(a), the blue curve is the raw data, the orange curve is the prediction result obtained using the LSTM without noise reduction, the green curve is the prediction result obtained using mean filter-LSTM, and the red curve is the prediction result obtained using the OTML prediction model. In Figure 8(b), the orange curve shows the prediction residuals when using the LSTM without filter, the green curve shows the predicted residuals when forecasting using mean filtering-LSTM, and the red curve shows the prediction residuals of the OTML prediction model. It can be seen that the prediction results of LSTM without filter have the largest error with the original data, and the prediction results obtained by mean filtering-LSTM model are improved, but the error between the predicted value and the true value is still large. In contrast, the OTML prediction model has the best prediction effect on the experimental data, which basically coincides completely with the original data.

Table 4 shows that RMSE, MAE, and MAPE are used as indicators to evaluate the prediction accuracy, and the prediction accuracy of the OTML prediction model, the mean filtering-LSTM

model and the LSTM without filter are evaluated, and the values of each performance index are as follows:

Table 4: Predictive performance metrics of different LSTM model

Model	RMSE	MAPE	MAE
Non-filter LSTM	0.51498	0.00566	0.41447
Median-filter LSTM	0.16758	0.00537	0.21838
OTML	0.01894	0.00531	0.09613

It can be seen from Table 4 that the prediction results of the LSTM network without filter has the largest index values, indicating that the prediction accuracy of the model is the worst. Compared with the LSTM network without filter, the indicators of the mean filtering-LSTM network are reduced, indicating that the prediction accuracy is improved. In contrast, the OTML predictive model has the smallest indicators, which verifies that the proposed predictive model can directly predict noisy data with high accuracy.

CONCLUSIONS

When it comes to network training and prediction based on the noisy data, improving the outlier-tolerance of the network is of great significance to improve the training performance and credibility of the results. In this paper, an OTML prediction model integrating outlier-tolerant filtering and multi-layer LSTM is built for outlier-tolerant prediction with high prediction accuracy of noisy data. The model integrates the Q-fault-tolerant filtering algorithm and the three-layer LSTM network, and realized the fault-tolerant prediction of the data containing random noisy, isolated/ patchy outliers. At the same time, the prediction effect of LSTM without filter and mean filtering-LSTM and the algorithm proposed in this paper on the same experimental data is compared and analyzed. The results showed that compared with the LSTM prediction results without filter, the accuracy of the prediction results of mean filtering-LSTM is improved, but the influence of noise and wild values cannot be completely avoided, and the OTML model proposed in this paper, whether in the face of isolated outliers or patchy outliers with a width less than the window radius, can minimize the influence of outliers and ensure the authenticity of the prediction results, which shows the feasibility and effectiveness of the proposed method. It can be expected that this method can expand and fuse more kinds of neural networks, such as radial basis neural networks (RBF), generative adversarial networks (GANs) and other types of attention mechanisms, to form hybrid neural network prediction models with higher prediction accuracy.

References

[1]. Karimanzira, D., Ritzau, L., & Emde, K. (2022). Catchment Area Multi-Streamflow Multiple Hours Ahead Forecast Based on Deep Learning. *Transactions on Engineering and Computing Sciences*, *10*(5), 15–29. https://doi.org/10.14738/tmlai.105.13049

[2]. namaswa, S., githiri, john., Mariita, N.., k'Orowe, M..., & Njiru, N. (2021). Application of Data-Driven Discovery Machine Learning Algorithms in Predicting Geothermal Reservoir Temperature from Geophysical Resistivity Method. *Transactions on Engineering and Computing Sciences*, *9*(4), 1–12. https://doi.org/10.14738/tmlai.94.10281

[3]. Zhang Lewen. Research on key technology of multi-component gas monitoring in heating furnace combustion process[D]. University of Science and Technology of China, 2022.DOI: 10.27517/d.cnki.gzkju.2022.000425.)

[4]. WANG Xin, WU Ji, LIU Chao, YANG Haiyan, DU Yanli, NIU Wensheng. Failure time series prediction based on LSTM recurrent neural network[J]. Journal of Beihang University,2018,44(04): 772-784.DOI:10.13700/j.bh.1001-5965.2017.0285.)

[5]. WANG Lipeng, LIANG Qi, ZOU Xianli, et al. Spectral noise reduction analysis method based on optimized moving average filtering[J]. Electrical Engineering,2018,19(11):43-45+54.)

[6]. YUE Haiyan, YAO Hong, ZHOU Zunsi. Comparison and simulation of filtering algorithms[J]. Wireless Internet Technology,2017(08):103-104.)

[7]. HU Shunbo. Fusion interpolation algorithm of B-spline filtering and GPVE in image registration[J]. Computer Applications,2011,31(08):2225-2228.)

[8]. J. Boudet, E. Lévêque, P. Borgnat, A. Cahuzac, M.C. Jacob. A Kalman filter adapted to the estimation of mean gradients in the large-eddy simulation of unsteady turbulent flows. Computers & Fluids, Volume 127, 2016, Pages 65-77, ISSN 0045-7930, https://doi.org/10.1016/j.compfluid.2015.12.006

[9]. CHEN Long. Prediction method of process variables driven by complex industrial data[D]. Dalian University of Technology, 2019.DOI: 10.26991/d.cnki.gdllu.2019.004119.)

[10]. Sheng Chunyang. Prediction method and application of secondary energy system of steel enterprises based on data[D]. Dalian University of Technology,2014.)

[11]. Wang Zhilan, Zhao Hongjie, Zhao Fan, et al. Establishment of remaining life prediction model of an inertial guide system based on convolutional neural network and filter fusion algorithm[J]. Aerospace Defense,2023,6(01):70-77.)

[12]. RONG Jing, LIU Lilong, WEN Hongyan, et al. Application of GM (1,1)-AR prediction model based on Kalman filter in deformation prediction[J]. Journal of Guilin University of Technology,2018,38(02):301-305.)

[13]. YI Lirong, WANG Shaoyu, YIN Lili, YANG Qing, GU Xin. Prediction of industrial sensor time series data based on multivariate LSTM[J]. Intelligent Computer and Application,2018,8(05):13-16.)

[14]. ZHU Feilin, HOU Tiantian, LIANG Yifan, REN Ruijie. Long-term runoff prediction model in reservoir based on VMD-LSTM[J/OL]. Hydroelectric Power:1-8[2023-05-25]. http://kns.cnki.net/kcms/detail/11.1845.TV.20230423.1719.004.html