Proceedings of the 5th International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/116/2025.20424

Research on Artificial Intelligence Methods for Feature Extraction and Prediction of Current Fluctuations

Zhaowen Huang[1,a,*]

[1]The University of Hong Kong, Hong Kong, China
a. huangzhaowen62@gmail.com
**Corresponding Author*

Abstract: Accurate prediction of current fluctuations in power systems is crucial for ensuring the safety and stability of grid operations. Traditional prediction methods face significant limitations when dealing with data complexity and dynamic changes. The introduction of artificial intelligence technologies offers a new perspective for current fluctuation predictions. Based on Long Short-Term Memory networks (LSTM), this study proposes an intelligent prediction method combining deep fusion and feature extraction of multi-source data. This method processes multi-dimensional information such as historical current data, weather conditions, and load demands to achieve high-accuracy and real-time predictions of current fluctuations. The results show that this method significantly improves prediction accuracy, enhances model adaptability, and provides reliable support for power dispatch and resource optimization, offering important reference value for the development of smart grids.

Keywords: Current fluctuations, Artificial Intelligence, Long Short-Term Memory Network, Feature extraction, Intelligent prediction

1. Introduction

The power system is a critical foundation for the operation of modern societal economics, and its stability directly affects industrial production, residential life, and economic development. Current fluctuations, as a common phenomenon in the operation of power systems, can lead to serious consequences such as voltage anomalies, equipment damage, or even large-scale blackouts. With the increasing integration of renewable energy, the volatility and complexity of power systems are further exacerbated. Traditional methods for predicting current fluctuations show clear deficiencies when facing large-scale, high-dimensional dynamic data, struggling to meet the needs for real-time and accurate predictions. The rise of artificial intelligence technology provides efficient tools for handling complex nonlinear data, particularly in the field of time series analysis. Long Short-Term Memory networks (LSTM), with their advantages in capturing long-term dependencies, are becoming a key means to address current fluctuation prediction issues. By incorporating deep learning techniques and combining the fusion and feature extraction of multi-source data, the accuracy and real-time performance of current fluctuation predictions have been significantly improved. The application of this technology not only optimizes power dispatch strategies but also effectively reduces resource waste and operational costs, providing technical support for the construction of smart grids.

DOI: 10.54254/2755-2721/116/2025.20424

2. Model Construction and Training

2.1. Model Selection

When making predictions for time series fluctuations in current, selecting the appropriate model is crucial. Long Short-Term Memory networks (LSTM) are a special type of recurrent neural network (RNN) designed to solve the vanishing gradient problem that traditional RNNs face when processing long sequence data. LSTMs introduce a gating mechanism that regulates the flow of information, effectively maintaining long-term dependencies while forgetting unnecessary data. This makes them perform exceptionally well in various prediction tasks, especially in applications like current fluctuation prediction that require capturing time-dependent relationships[1]. The core feature of LSTM is its gating mechanism, which includes forget gates, input gates, and output gates. These gates help the network learn when to retain or forget information, effectively solving long-term dependency issues.

Forget Gate:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

Here, $f_t$the output of the forget gate at time step t is denoted, where$\sigma$is the sigmoid activation function,W_f is the weight matrix for the forget gate, $b_f$is the bias term,h_{t-1} is the hidden state from the previous time step, and $x_t$is the input at the current time step.

Input Gate and Cell State Update:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_i] + b_i)$$

$$\tilde{C}_t = tanh(W_C \cdot [h_{t-1}, x_t] + b_c)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

Here, $i_t$is the output of the input gate,$\tilde{C}_t$is the new candidate value,C_t is the updated cell state,$*$ representing element-wise multiplication.

Output Gate:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_c)$$

$$h_t = o_t * tanh(C_t)$$

Here,$o_t$is the output of the output gate, $h_t$which is the current time step's output or hidden state.

LSTM is capable of handling data sequences of varying lengths, has strong adaptability, and is particularly effective for the current fluctuation data in power systems. It can accurately capture the dynamic changes in fluctuations, which is crucial for prediction accuracy.

2.2. Model Construction

When building the LSTM model, the focus is on the design of the model's structure and parameter settings to ensure high performance and accuracy in predictions.

The model architecture includes the following parts:

Input Layer: Designed according to the dimensionality of the input features, typically including historical current data, meteorological conditions, and load demands.

LSTM Layers: Multiple LSTM layers can be included to increase the depth of the model, allowing it to handle more complex data features. The number of neurons per layer and the number of layers are determined based on the complexity of the actual data and the length of dependencies that need to be captured[2].

(3) Output Layer: The output layer structure is designed according to the requirements of the prediction task, such as predicting current fluctuations over a future period.

DOI: 10.54254/2755-2721/116/2025.20424

Key parameter settings are as follows:

Learning Rate: Typically set between 0.001 and 0.01, adjusted based on the model's performance in the early stages of training.

Batch Size: Affects the stability and convergence speed of model training, generally set at 32 or 64.

Epochs: Determined based on the model's performance on the training and validation sets, often setting an early stopping strategy to prevent overfitting.

(4) Optimizer: The Adam optimizer is chosen for its combination of advantages from Adagrad and RMSprop optimizers, suitable for handling optimization issues in unstable deep learning networks.

(5) Regularization Techniques: Such as Dropout or L2 regularization, used to reduce overfitting and improve the model's generalization ability.

Through this carefully designed LSTM model, the prediction accuracy of current fluctuations can be effectively improved, providing scientific decision support for the stable operation of power systems. This not only enhances the reliability of the power grid but also lays a solid foundation for the development of future smart grids.

2.3. Model Training

Model training is a crucial step to ensure that the prediction model can effectively and accurately reflect real-world conditions[3]. In the training process of Long Short-Term Memory networks (LSTM), selecting the appropriate loss function and optimizing hyperparameters are important aspects for achieving efficient training and improving model performance.

In the current fluctuation prediction model, the commonly used loss function is the Mean Squared Error (MSE), which is a standard method for measuring the difference between predicted values and actual values. The formula for MSE is:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

where $n$is the number of samples, y_i are the actual values, and $\hat{y}_i$are the predicted values. MSE helps to amplify larger errors, making the model more focused on reducing the occurrence of large errors during training, which is particularly important for predicting current fluctuations, as sudden large fluctuations in current can significantly impact the stability of the power grid.

The steps for tuning hyperparameters are as follows:

Learning Rate: The learning rate is a key parameter that affects the efficiency and convergence speed of model training. If the learning rate is too high, it may cause the model parameters to update too aggressively during training, making it difficult to converge; if the learning rate is too low, the training speed is slow, and it is easy to get stuck in local optima. Therefore, optimizers like Adam, which have the capability to adaptively adjust the learning rate based on changes in the gradient, are commonly used to automatically adjust the learning rate during training.

Batch Size: The batch size determines the number of data samples used to compute gradient descent in each iteration. Smaller batches can enhance the model's generalization ability but may make the training process more unstable[4]. Larger batches can stabilize the training process but increase memory consumption and may lead to overfitting. Therefore, choosing the appropriate batch size has a direct impact on model performance.

(3) Epochs: The number of epochs needs to be sufficient to ensure that the model can fully learn the features in the training data. At the same time, too many epochs might lead to overfitting. Typically, early stopping techniques are combined to terminate training when there is no further improvement on the validation set, to prevent overfitting.

Through these strategies, the LSTM model can be effectively trained to achieve high accuracy and stability in predicting current fluctuations. Additionally, timely model evaluation and parameter adjustment are key to ensuring long-term stable operation.

3. Prediction Implementation and Results Analysis

3.1. Deployment of Real-Time Prediction System

Deploying a real-time prediction system typically includes integrating the trained model into the existing power system monitoring architecture. The specific steps are as follows: System architecture and integration are as follows:

Data Collection: Deploy a data collection system to collect real-time sensor data such as current, voltage, temperature, and wind speed from the power system. The data collection frequency is set to once per minute to ensure sufficient data density to capture rapid changes in current.

Data Preprocessing: Real-time preprocessing includes data cleaning (filtering out anomalies), normalization (scaling data between 0 and 1), and feature engineering (such as rolling window averages). These steps are performed in real-time using automated scripts before the data enters the model.

Model Deployment: Deploy the LSTM model on servers or in a cloud environment. The model is deployed in a high-performance computing environment, such as AWS EC2 or Azure VM, to support large-scale parallel processing and real-time responses.

Model parameters and configuration are as follows:

LSTM Configuration: The model has three LSTM layers, each with 128 units. This configuration is sufficient to handle complex time series data and capture long-term dependencies.

Real-Time Data Interface: Real-time data flows into the prediction system via an API from the data collection system. The API uses secure protocols (such as HTTPS) to ensure the security of data transmission.

Real-time prediction execution steps are as follows:

Prediction Cycle: The system is set to run a prediction every 5 minutes, predicting current fluctuations over the next 60 minutes.

Output Handling: After the prediction results are denormalized, they are directly sent to the monitoring screens at the power dispatch center and notified to relevant personnel via email and SMS.

3.2. Results Analysis and Validation

3.2.1. Data Collection and Preparation

The test dataset collected historical current data from the past year as the test dataset, which was systematically recorded and stored in a highly secure data warehouse, totaling over 520,000 records, covering the full year from January to December 2023.

3.2.2. Performance Evaluation Metrics

In the evaluation of prediction model performance, Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) are three commonly used metrics. MSE is one of the standard methods for measuring the difference between predicted values and actual values. It calculates the average of the squared differences between all predicted and actual values[5]. RMSE is the square root of the mean squared error, providing an error size assessment in the same units as the original data. MAPE measures the average proportion of prediction value errors to actual values.

3.2.3. Evaluation Results

The results of the model evaluation over different periods are shown in Table 1 as follows:

Table 1: Results of Model Evaluation Across Different Time Periods

Time Period	MSE	RMSE	MAPE
February 2023	0.0105	0.1025	2.98%
April 2023	0.0096	0.0980	2.75%
June 2023	0.0090	0.0949	2.60%
August 2023	0.0084	0.0916	2.45%
October 2023	0.0081	0.0900	2.30%
December 2023	0.0078	0.0883	2.20%

The table shows that from February to December, the accuracy of the model gradually improved, indicating that through continuous training and fine-tuning of the model, its predictive ability was significantly enhanced. Especially in handling the dynamic and complex issue of current fluctuations, the model was able to adapt and predict accurately. The monthly reduction in MSE and RMSE reflects the model's gradually enhanced ability to accurately capture the characteristics of current fluctuations, and the continuous decline in MAPE indicates that the proportion of errors in model predictions is decreasing. Particularly, the continuous maintenance of MAPE below 5% demonstrates the model's high practical value and reliability in real applications.

4. Conclusion

By applying the Long Short-Term Memory (LSTM) model to predict current fluctuations, this study has demonstrated the potential application of deep learning technology in actual power system management. The deployment of the LSTM model not only enhanced real-time monitoring capabilities for current fluctuations but also improved the efficiency of early warning systems for power systems in responding to emergencies. Furthermore, the model's continuous data learning and self-adjustment have shown excellent adaptability and accuracy, providing valuable experience and references for the development of future smart grid technologies.

References

[1] *Zank P G ,Zhao L L ,Adhikari L , et al. Characterization of Turbulent Fluctuations in the Sub-Alfvénic Solar Wind [J]. The Astrophysical Journal, 2024, 966 (1):*

[2] *Hussain S ,Madathodika A ,Rachid E , et al. Rule-Based Non-Intrusive Load Monitoring Using Steady-State Current Waveform Features. [J]. Sensors (Basel, Switzerland), 2023, 23 (15):*

[3] *Rongke S ,Huichang G . Leakage Protection of Low Voltage TN-C System Based on Characteristic Currents Amplitude Ratio [J]. Journal of Physics: Conference Series, 2023, 2477 (1):*

[4] *Li C ,Qiang H ,Mi Z , et al. Characteristics of Continuing Current Waveforms and M-Component Parameters in Triggered Lightning [J]. Journal of Geophysical Research: Atmospheres, 2022, 127 (23):*

[5] *Giuseppe C ,Paola M D ,Tommaso A , et al. Multiscale Features of Magnetic Field Fluctuations and Field-Aligned Currents in the Polar Ionosphere: A Case Study [J]. Universe, 2022, 8 (12): 610-610.*