A Study on an Artificial Intelligence Prediction Method for Current Fluctuations Based on Multivariate Analysis

Zhaozhang Huang^{1,a,*}

¹Ulink High School of Suzhou industrial Park a.changchang819@126.com *Corresponding Author

Abstract: With the increasing complexity of power systems and higher demands for stability, accurately predicting current fluctuations is crucial. This study proposes an artificial intelligence prediction method combined with multivariate analysis, utilizing multidimensional information such as historical current data, new energy generation output, and meteorological conditions, and modeling through Long Short-Term Memory (LSTM) networks. This method optimizes the data processing process and improves the model's ability to adapt to dynamic changes in data. The results show that the proposed method effectively enhances the prediction accuracy of current fluctuations, providing practical value for grid operation and maintenance as well as fault warnings.

Keywords: current fluctuations, artificial intelligence, multivariate analysis, prediction model, Long Short-Term Memory (LSTM)

1. Introduction

Current fluctuation prediction plays a crucial role in maintaining the stability and safe operation of power systems. Traditional prediction methods often rely on single data sources or simplified mathematical models, which are insufficient when handling the dynamics of complex power systems. With the development of big data technology and advancements in artificial intelligence algorithms, efficiently integrating and utilizing current data, new energy generation data, and meteorological information has become key to improving prediction accuracy. This study implements advanced multivariate analysis and deep learning techniques, particularly Long Short-Term Memory (LSTM) networks, which significantly improve the accuracy of current fluctuation prediction while enhancing the model's adaptability to changes in the power system. This is of great significance for real-time grid monitoring and emergency response strategy formulation, helping achieve more precise and economical grid management.

2. Data Preparation and Preprocessing

2.1. Data Collection

The focus of data collection is on acquiring various types of data directly related to current fluctuations, including historical current data, new energy generation data, and meteorological data.

These data are sourced from official records of the National Grid and local meteorological stations to ensure accuracy and reliability.

Historical Current Data: Core data for predicting current fluctuations, including hourly current readings collected from the power system from 2018 to 2023. Historical data helps the model capture seasonal variations and long-term trends in current fluctuations[1].

New Energy Generation Data: With the widespread adoption of new energy technologies, solar and wind power generation significantly impact the grid. This study collects hourly new energy generation data for the same period, reflecting current variations due to renewable energy integration[2].

Meteorological Data: Meteorological conditions directly affect the efficiency of new energy generation and power load demand. This study includes hourly meteorological indicators such as temperature, humidity, and wind speed.

2.2. Data Preprocessing

Data preprocessing is a critical step in data analysis as it directly impacts the performance and accuracy of the prediction model. It includes two main stages: data cleaning and normalization.

Data Cleaning: [3]The collected data undergoes quality checks to eliminate errors or anomalies. For example, outliers are identified and handled using statistical analysis methods such as box plots and the 3σ principle. Missing data points are filled using linear interpolation or time series forecasting methods based on adjacent data. Duplicate records are also removed to ensure the uniqueness and consistency of the dataset.[4]

Normalization: Current data, new energy generation data, and meteorological data differ in units and dimensions. Directly using these data can lead to bias in the model training process. Normalization is applied to scale all numerical features to a common range, typically between 0 and 1. The normalization formula is as follows:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} x' = \frac{x - \min(x)}{\max(x) - \min(x)}$$

Here, x is the original data point, min(x) and max(x) are the minimum and maximum values in the dataset. This ensures balanced contributions of different features to the model, improving training efficiency and prediction accuracy.[5]

3. Feature Selection and Extraction

3.1. Feature Selection Methods

Feature selection is a key step in constructing a prediction model. Proper feature selection improves prediction accuracy while significantly reducing the computational resources required for model training. This study uses the Pearson correlation coefficient method, a statistical approach, for initial feature screening.

The Pearson correlation coefficient measures the linear correlation between two variables, with values ranging from -1 to 1. A coefficient of 1 indicates a perfect positive correlation, -1 indicates a perfect negative correlation, and 0 indicates no linear correlation. The study calculates the Pearson correlation coefficients between historical current fluctuation data and other potential predictors (e.g., new energy generation data and various meteorological parameters). This analysis quickly identifies variables with significant linear relationships to current fluctuations. The formula for the Pearson correlation coefficient is:

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}$$

Here, x_i and y_i are the values of variables, X and Y at the i-th sample point, \bar{x} and \bar{y} are the sample means of the respective variables.

Through this calculation, highly correlated features, such as new energy generation under specific meteorological conditions and power load demand at particular time points, are identified.

3.2. Implementation of Feature Engineering

After the initial feature selection, machine learning algorithms are further applied to optimize the feature selection and extraction process. Machine learning is particularly effective in identifying and leveraging information in complex patterns, especially when dealing with nonlinear relationships and interaction effects.

Principal Component Analysis (PCA): PCA is a commonly used feature extraction method that reduces redundant information in the dataset through dimensionality reduction while retaining the most critical variables. In this study, PCA is used to further compress and optimize the selected feature set, particularly for high-dimensional meteorological and power load data. By transforming multiple correlated variables into a few independent principal components, PCA explains most of the data variability while retaining the most important information for the prediction model.

4. Model Construction and Training

4.1. Model Selection and Construction

To accurately predict current fluctuations, this study selects the Long Short-Term Memory (LSTM) network as the prediction model. LSTM is particularly suitable for handling time-series-related problems due to its ability to effectively manage long-term dependencies, which is critical for predicting current fluctuations.

The reasons for selecting LSTM are as follows:

(1)Handling Long-Term Dependencies: Traditional Recurrent Neural Networks (RNNs) struggle to capture dependencies in long time series, while LSTM's gating mechanism effectively solves this problem.

(2)Adapting to Complex Variability: Power system data typically exhibit high nonlinearity and complexity. LSTM, with its internal state update mechanism, adapts to such complexities and provides stable and accurate predictions.

The model structure consists of the following components:

(1)Input Layer: Receives specific input variables based on feature selection results, such as historical current data, new energy generation outputs, and critical meteorological parameters.

(2)LSTM Layer: Multiple LSTM layers are constructed to enhance the model's ability to capture data features. Each LSTM unit consists of a forget gate, input gate, and output gate, which collectively determine the flow and retention of information.

(3)Fully Connected Layer: Outputs from multiple LSTM layers are connected to a fully connected layer, which integrates learned features and maps them to the final output.

(4)Output Layer: Produces the final prediction output, usually adjusted with an activation function to match the actual range of the predicted values.

4.2. Model Training

Model training is a key step in achieving efficient predictions. This study adopts a series of strategies to optimize the training process. Firstly, the historical dataset is divided into training, validation, and

test sets. The training set is used for model training, the validation set for tuning parameters and preventing overfitting, and the test set for evaluating the model's final performance.

To quantify prediction error, the Mean Squared Error (MSE) is selected as the loss function. MSE, the mean of the squared differences between predicted and actual values, is suitable for regression problems and effectively reflects prediction accuracy in current fluctuation prediction:

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

Where, Y_i is the actual value, \hat{Y}_i is the predicted value, and n is the total number of samples.

For optimization, the Adam optimizer is employed, an adaptive learning rate algorithm combining the advantages of AdaGrad and RMSProp. Adam automatically adjusts the learning rate based on the problem's characteristics, making it suitable for handling large-scale and complex datasets.

During the training process, network parameters are iteratively adjusted to minimize the loss function. Early stopping is used to monitor performance on the validation set, preventing overfitting and ensuring good generalization ability.

5. Real-Time Prediction and Decision Support

5.1. Implementation of Real-Time Prediction System

The real-time prediction system deployed in this study is based on the advanced LSTM network. This system not only processes real-time incoming data but also outputs prediction results instantly, supporting power grid operation and maintenance.

The core of the system is a real-time data processing and prediction module, which continuously receives current data, new energy generation data, and meteorological data from the power system. These data are transmitted in real time through predefined APIs, ensuring timeliness and completeness. The received data first undergo a preprocessing phase, including cleaning and normalization, to ensure input data quality meets the model's requirements.

The preprocessed data are then input into the trained LSTM model for real-time prediction. The optimized real-time performance of the LSTM model enables the system to complete complex prediction calculations within seconds, ensuring rapid response.

5.2. Application of Prediction Results and Decision Support

The real-time prediction results generated by the system must not only be accurate but also provide practical value for grid management and maintenance decisions. The comparison between predicted and actual current fluctuation values at specific time points in a power system is shown in Table 1:

Table 1: Comparison of Predicted and Actual Current Fluctuation Values at Specific Time Points in a Power System

Time Point	Predicted Current Fluctuation (A)	Actual Current Fluctuation (A)	Prediction Error (A)
2023-12-01 00:00	3.5 A	3.6 A	0.1 A
2023-12-01 01:00	2.8 A	2.7 A	0.1 A
2023-12-01 02:00	4.2 A	4.1 A	0.1 A
2023-12-01 03:00	5.0 A	4.8 A	0.2 A
2023-12-01 04:00	3.3 A	3.3 A	0.0 A

The accuracy of the prediction results directly affects the effectiveness of decision support. In practical applications, the predicted data can be used for various decision support activities:

Risk Assessment: The system evaluates the risk level of current fluctuations in real time based on preset safety thresholds and issues automatic warnings when the predicted values exceed these thresholds, allowing maintenance personnel to adjust grid operations or troubleshoot faults promptly.

Load Distribution Optimization: By predicting current fluctuations, grid dispatchers can adjust load distribution in advance, optimize power resource utilization, and reduce energy waste.

Maintenance Planning: Long-term current fluctuation prediction data helps grid companies develop more effective maintenance plans and budget allocations.

By deploying and operating the real-time prediction system, the safety and economic efficiency of grid operations are significantly enhanced, ensuring stable and reliable power supply.

6. Conclusion

By integrating multivariate analysis and Long Short-Term Memory (LSTM) networks, this study develops a model capable of effectively predicting current fluctuations, providing strong technical support for real-time monitoring and intelligent management of power systems. The successful deployment of the real-time prediction system optimizes grid operation strategies, enhances the accuracy and timeliness of warnings, and further ensures the stability of power supply and economic grid operation. Moreover, the model demonstrates high efficiency and adaptability in handling large-scale and complex datasets, proving the potential of deep learning technologies in solving traditional industry challenges. The outcomes of this study not only advance power system analysis techniques but also provide a reference model for other industries, showcasing how machine learning can address and predict complex variable relationships in big data environments.

References

- [1] Coghi F. Current fluctuations of a self-interacting diffusion on a ring [J]. Journal of Physics A: Mathematical and Theoretical, 2025, 58 (1): 015002-015002.
- [2] Berlioz T, Venturelli D, Grabsch A, et al. Current fluctuations in the symmetric exclusion process beyond the onedimensional geometry [J]. Journal of Statistical Mechanics: Theory and Experiment, 2024, 2024 (11): 113208-113208.
- [3] Zargari M H ,Talavat V ,Bolandi G T . A fast, simple and local protection scheme for fault detection and classification during power swings based on differential current component [J]. International Journal of Electrical Power and Energy Systems, 2024, 163 110369-110369.
- [4] Riazantseva O M, Treves V T, Khabarova O, et al. Linking Turbulent Interplanetary Magnetic Field Fluctuations and Current Sheets [J]. Universe, 2024, 10 (11): 417-417.
- [5] Lim N ,Hong S ,Jung J , et al. Enhancing mixed gas discrimination in e-nose system: Sparse recurrent neural networks using transient current fluctuation of SMO array sensor [J]. Journal of Industrial Information Integration, 2024, 42 100715-100715.