ISSN 1330-3651 (Print), ISSN 1848-6339 (Online)

https://doi.org/10.17559/TV-20241026002094
Original scientific paper

# Integrating Meteorological Data and Bidirectional LSTM Models for Vulnerability Detection and Security Strategy Development in Power Energy Systems

Hongxia WANG, Jie XU*, Yang YANG, Meng LI

**Abstract:** The increasingly complex power energy system, coupled with the increasing frequency of extreme weather events, requires powerful models for vulnerability detection and security strategy development. This study proposes a novel approach that combines meteorological data with a bidirectional Long Short Term Memory (LSTM) model enhanced by attention mechanisms. This model effectively captures the dynamic interaction between time dependence, meteorological factors, and system vulnerabilities. The key performance indicators, including accuracy, precision, and recall, demonstrate that the model has excellent predictive ability compared to traditional methods. The results show that the average accuracy of the bidirectional LSTM model for vulnerability classification of power energy systems is as high as 97.3%, and the recall fluctuation in different testing environments is only 0.2%, which is superior to traditional techniques such as Unidirectional LSTM and Support Vector Machine (SVM). Numerical experiments have verified the robustness of the model in different scenarios, while case studies have demonstrated its practical applicability in identifying key vulnerabilities and notifying targeted security measures. The unique combination of meteorological data and power system operation data enables the bidirectional LSTM model to accurately analyze complex environmental impacts and effectively identify potential vulnerabilities. This study contributes to the development of energy system prediction and analysis, providing a scalable and adaptable framework to enhance the system's ability to resist meteorological impacts. Future work will focus on integrating real-time monitoring systems and extending the applicability of models to a wider range of energy infrastructure environments.

**Keywords:** attention mechanism; bidirectional long short-term memory; meteorological dana; power energy system; security protection strategies; vulnerability analysis

## 1 INTRODUCTION

Power energy systems are increasingly vulnerable to extreme weather events, which have caused significant disruptions, economic losses, and safety risks globally. Recent reports indicate that weather-related power outages in 2022 alone accounted for over $150 billion in damages [1]. Traditional vulnerability detection methods, often reliant on static or linear models, fail to capture the dynamic and nonlinear interactions between meteorological factors and system vulnerabilities [2, 3]. These limitations underscore the need for advanced, data-driven approaches that can integrate diverse meteorological datasets and provide real-time risk assessments. This study addresses these challenges by proposing a novel bidirectional Long Short-Term Memory (LSTM) model with an attention mechanism, which dynamically captures temporal dependencies and highlights critical meteorological features affecting power energy system resilience."In this part, you need to refine the problem statement and contextulize the proposed method here. You also need to add a clear and concise statement of the research objectives, such as: "This study aims to develop a scalable and adaptive risk prediction model that integrates meteorological data with advanced machine learning techniques to improve power energy system resilience.

The primary research objectives of this study are: (1) To develop a robust, data-driven framework that integrates operational and meteorological data for real-time vulnerability detection. (2) To improve classification accuracy and stability compared to traditional models such as unidirectional LSTM, Support Vector Machines (SVM), and Decision Trees (DT). (3) To provide actionable insights for targeted security strategies, enhancing system resilience against extreme weather conditions.

By building on and addressing gaps in prior studies, this work establishes a more comprehensive approach to vulnerability analysis in power energy systems. The proposed bidirectional LSTM model demonstrates superior performance by simultaneously capturing temporal dependencies, multivariate interactions, and the influence of meteorological conditions, providing a significant step forward in predictive analysis and system protection.

## 2 RELATIVE WORKS

### 2.1 Impact of Meteorology on the Power System

The impact of meteorology on the power system is mainly reflected in three aspects: power generation, transmission and distribution, and electricity demand. Meteorological conditions directly affect the power generation capacity of renewable energy sources such as hydropower, wind power, and photovoltaics. For example, precipitation determines hydropower output, while wind speed and sunshine affect the efficiency of wind and photovoltaic power generation. High temperatures may reduce transmission efficiency and cause overheating of power lines, while low temperatures may lead to line icing and disconnection. Extreme weather conditions such as thunderstorms and typhoons can also damage transmission and distribution equipment, resulting in widespread power outages. In terms of electricity demand, temperature changes can lead to significant fluctuations in air conditioning and heating electricity consumption, and extreme weather can also increase the difficulty of load regulation [4]. In addition, meteorological disasters and changes have profound impacts on power grid scheduling, equipment planning, and market prices, which need to be addressed through meteorological forecasting, equipment reinforcement, and energy storage technology.

### 2.2 Research on Power Energy Systems

The stable operation of the power energy system is crucial for ensuring the normal operation of the social economy. In-depth analysis of system vulnerabilities helps to develop effective security protection strategies and ensure the continuous and stable operation of the system. The complexity of the power system is constantly increasing, and the power system is facing various potential vulnerabilities and security threats. A comprehensive analysis of vulnerabilities in the power energy system can help ensure the stable operation of the power system. Sahoo and other scholars have designed a power grid system integrity protection technology that includes offline computing and real-time local information settings, which determines the load to be dropped based on the sensitivity calculation of relay operation margin [5]. Bedi Guneet used the Internet of Things to digitize the power ecosystem, which can improve asset visibility, optimize distributed generation management, eliminate energy waste achieve savings, and conduct potential system vulnerability monitoring [6]. Venkatanagaraju et al. developed a current supervised detection index based on matrix pencil method, which adaptively detects power system faults after signal reconstruction [7]. Regular detection of vulnerabilities in the power system can identify potential system vulnerabilities and security risks, prevent potential security risks in advance, and ensure that the power system is not affected by malicious attacks or natural disasters [8-9]. Previous research has mainly focused on the technical aspects of power energy systems, neglecting the impact of meteorological data on the system.

### 2.3 Research on Meteorological Data Analysis Methods

Analyzing meteorological data helps to understand the operational mechanisms of climate systems and the formation and variation patterns of climate patterns. Many people have predicted the operational status of a system by comprehensively analyzing meteorological data, in order to develop targeted safety protection strategies. Climate factors such as temperature and humidity can seriously affect the operation status of power system equipment. Predictive analysis using meteorological data can help the power system take proactive measures to adapt to changing conditions [10, 11]. Kumar et al. used machine learning techniques to construct an automatic analysis framework for meteorological pollution information, exploratory data analysis, and in-depth understanding of various hidden patterns in the dataset [12]. Scholars such as Poddar have developed an automatic feature extraction technique using deep learning methods to learn and analyze meteorological and remote sensing data to predict corresponding crop growth [13]. Rahman Obaidur's research indicated that voltage imbalance in distribution systems is related to meteorological data such as temperature. Superconducting magnetic energy storage through high-temperature superconducting coils can be integrated into other commercial battery systems, forming a hybrid energy storage system that can alleviate the potential problems brought about by large-scale penetration of distributed power generation [14]. Comprehensive analysis of meteorological data and power system operation data can effectively detect system anomalies and provide targeted safety precautions. However, there is a lack of effective models for comprehensive analysis of meteorological data.

On the basis of previous work, a bidirectional LSTM model with an attention mechanism was proposed, which significantly improves the ability to capture temporal dependencies and feature correlations compared to traditional methods. Traditional models often only handle unidirectional time series and cannot comprehensively analyze the complex relationship between power system operation data and meteorological data. This study uses bidirectional LSTM to simultaneously capture the dependency relationship between time steps before and after and introduces an attention mechanism to dynamically focus on the most relevant part of the input sequence to the current output, greatly improving the model's classification performance for power system vulnerabilities.

## 3 METHODS FOR ADDRESSING VULNERABILITIES IN POWER ENERGY SYSTEMS

### 3.1 Collection of Operational and Meteorological Data for Power Energy Systems

As one of the important infrastructures in modern society, the power energy system is facing increasingly complex and diverse threats. The power energy system is the foundation for maintaining economic development and social stability, and any attack or vulnerability to the power system can lead to serious consequences [15, 16].

The level of informatization and intelligence in the power energy system is constantly improving, and the vulnerability of the system is correspondingly increasing [17, 18]. By making reasonable use of the data generated by the power energy system, potential vulnerabilities can be identified and corrected, ensuring that the system operates in an efficient, stable, and safe state.

The meteorological environment in which a power energy system is located is strongly linked to the operation and performance of the power system. Wind and solar energy are common renewable energy sources, and the energy generated by the power system is influenced by climate conditions. Severe meteorological disasters may lead to power system interruptions, equipment damage, or insufficient energy supply, and factors such as temperature and humidity can affect the performance of power equipment.

In order to effectively analyze vulnerabilities in the power energy system, meteorological data and operational data of the power energy system are comprehensively analyzed. Meteorological data is collected through various sensors deployed in meteorological stations, including thermometers, hygrometers, anemometers, wind direction sensors, barometers, etc. The operational status of the power energy system is monitored through sensor networks in the power energy system.

The collected meteorological data information is described in Tab. 1.

In Tab. 1, the collected meteorological data information is described, which includes temperature, humidity, wind speed, and pressure information. All collected meteorological information includes a timestamp. Analyzing meteorological data can help predict future meteorological conditions and predict the power generation and load that affect the power energy system.

The collected operational data of the power energy system is described in Tab. 2.

In Tab. 2, the operating data of the power energy system is described. The collected data includes power generation, voltage, current, frequency, and load information, and also includes time labels. By monitoring the operational data of the power energy system, potential signs of failure can be identified. By identifying problems early and taking preventive measures to prevent power energy system downtime due to faults, the reliability and stability of the system can be improved.

**Table 1** Collected meteorological data information

| Timestamp | Temperature / °C | Humidity / % | Wind speed / m/s | Air pressure / HPa |
|---|---|---|---|---|
| 2023-01-01 12:00 | 20 | 60 | 4.2 | 1015 |
| 2023-01-02 12:00 | 24 | 56 | 5.3 | 1013 |
| 2023-01-03 12:00 | 23 | 49 | 6.8 | 1015 |
| 2023-01-04 12:00 | 26 | 56 | 2.9 | 1014 |
| 2023-01-05 12:00 | 22 | 53 | 3.9 | 1014 |
| 2023-01-06 12:00 | 17 | 54 | 4.5 | 1015 |
| 2023-01-07 12:00 | 24 | 58 | 4.4 | 1015 |
| 2023-01-08 12:00 | 24 | 54 | 4.9 | 1013 |
| 2023-01-09 12:00 | 27 | 53 | 5.3 | 1013 |
| 2023-01-10 12:00 | 23 | 59 | 3.9 | 1015 |

**Table 2** Operating data of power energy system

| Timestamp | Power generation / kWh | Voltage / V | Current / A | Frequency / Hz | Load / kW |
|---|---|---|---|---|---|
| 2023-01-01 12:00 | 1500 | 230 | 50 | 60 | 100 |
| 2023-01-02 12:00 | 1490 | 228 | 52 | 62 | 98 |
| 2023-01-03 12:00 | 1450 | 239 | 49 | 62 | 79 |
| 2023-01-04 12:00 | 1490 | 231 | 52 | 61 | 120 |
| 2023-01-05 12:00 | 1390 | 230 | 53 | 59 | 119 |
| 2023-01-06 12:00 | 1550 | 229 | 54 | 58 | 112 |
| 2023-01-07 12:00 | 1620 | 230 | 51 | 57 | 113 |
| 2023-01-08 12:00 | 1780 | 228 | 50 | 59 | 109 |
| 2023-01-09 12:00 | 1670 | 229 | 52 | 61 | 98 |
| 2023-01-10 12:00 | 1560 | 230 | 50 | 60 | 97 |

The structural model of the power energy system is illustrated in Fig. 1.

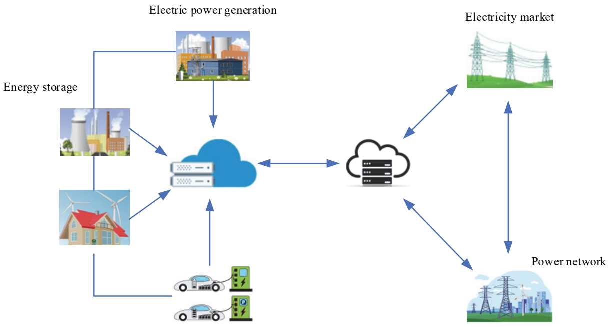


**Figure 1** The main structure of the power energy system

In Fig. 1, the structural model of the power energy system is described. Electricity is provided through power generation devices, which are applied through storage and coordinated optimization. By using a large centralized control platform to transmit electricity to the electricity market and large power networks, stable and reliable electricity is provided.

## 3.2 Data Preprocessing

The data collected from different sensors may have missing or abnormal values, and noise data is inevitably generated during the data collection and transmission process. To ensure the quality and availability of data, the collected meteorological data and power energy system operation data are preprocessed.

The content of data preprocessing includes data cleaning, time series processing, data denoising, data standardization, and extracting meaningful data features through feature engineering. The goal of data cleaning is to ensure the quality of raw data and detect abnormal data, missing item data, and duplicate data records through statistical methods [19, 20].

For the detected abnormal data and missing item data, the mean of domain data is used for replacement and filling. For the detected duplicate data records, the duplicate data information needs to be deleted and processed.

The collected meteorological data and power energy system operation data both contain timestamp attributes. Converting the timestamp to date time format and creating a time series index is more conducive to analyzing the time dependency relationship of the data.

The sensor collects data and the data transmission process generates noise data, which is denoised using Gaussian filtering. Gaussian filtering performs weighted averaging on data to suppress noise and smooth signals. The formula for the Gaussian function is expressed as:

$$g\left(x\right)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{1}$$

Data standardization is the process of converting data of different scales and ranges into data with a unified scale and range. The *z*-score normalization method is adopted to eliminate the influence of different data scales, and the standardized formula is expressed as:

$$z=\frac{\left(x-u\right)}{\sigma} \tag{2}$$

Eq. (2), $u$ represents the mean of the original data.

To analyze vulnerabilities in the power energy system more effectively, meaningful feature information is extracted through feature engineering. The mean is the average of a set of data, representing the central trend of the data, and is used to measure the center position of the data. The data mean is calculated by the formula:

$$Mean = \frac{\sum_{i=1}^{n} X_i}{n} \tag{3}$$

Variance measures the degree of dispersion of data, and it can help identify the variability of data. The formula for calculating variance is expressed as:

$$V = \frac{\sum_{i=1}^{n} \left( X_i - Mean \right)^2}{n} \tag{4}$$

Deviation measures the asymmetry of data distribution, with positive deviation indicating right deviation and negative deviation indicating left deviation. The calculation formula for data deviation is expressed as:

$$S = \frac{\sum_{i=1}^{n} \left( X_i - Mean \right)^3}{n \times V^{3/2}} \tag{5}$$

The collected data is converted into a data format that can be analyzed by LSTM models. The input requirement for the LSTM model is a three-dimensional array, including the count of samples, time steps, and the count of features. The input three-dimensional array form is:

[
[
$[k_{11}, k_{12}, ..., k_{1n}]$ <- time step feature vector
], <- the first time series sample
[
$[k_{21}, k_{22},..., k_{2n}]$ <- time step feature vector
], <- The second time series sample
...
[
$[k_{t1}, k_{t2}, ..., k_{tn}]$ <- time step feature vector
] <- The *t*-th time series sample
]

To label vulnerabilities for each time series data, the main types of vulnerabilities in the power energy system include: physical security vulnerabilities, network security vulnerabilities, software security vulnerabilities, data security vulnerabilities, Internet of Things (IoT) device vulnerabilities, and remote access vulnerabilities.

Time series data is labeled as vulnerabilities in the power energy system. The content of the data label is displayed in Tab. 3.

In Tab. 3, the content of the data labels is described. Six types of vulnerabilities and normal temporal data are statistically analyzed. The impact of different types of vulnerabilities on power energy systems varies, and network vulnerabilities may gain unauthorized access to the system, thereby manipulating or damaging it. Data security vulnerabilities may lead to the leakage of sensitive information in the system. In actual data preprocessing, the first step is to improve data quality by detecting and processing missing values, outliers, and duplicate records. Fill in missing values with mean and delete duplicate data. Secondly, convert the timestamp of the data into a time series index to capture temporal dependencies. Then, the Gaussian filtering method is used to smooth the noise in the sensor data and reduce interference. Finally, the z-score standardization method is used to uniformly scale the data, eliminate the influence of different dimensions, and extract key features such as mean and variance of the data through feature engineering, providing high-quality input data for the model.

**Table 3** The content of data labels for time series data

| Type | Number of time series data samples | Percentage |
|---|---|---|
| Physical security vulnerabilities | 4440 | 15.9% |
| Network security vulnerabilities | 4560 | 16.3% |
| Software security vulnerabilities | 4200 | 15.0% |
| Data security vulnerabilities | 4400 | 15.7% |
| Internet of Things device vulnerabilities | 3800 | 13.5% |
| Remote access vulnerability | 4000 | 14.3% |
| Normal | 2600 | 9.3% |

### 3.3 Building Bidirectional LSTM Models

LSTM network is a deep learning model used for processing sequence data, which can effectively solve the problem of vanishing or exploding gradients when processing long sequences, thereby better capturing long-term dependencies in sequences [21, 22]. LSTM introduces a cellular state that can transmit information in long sequences, allowing the network to selectively forget or store information, and can handle long-term dependencies [23, 24].

The operational data and meteorological data of the power energy system both have temporal attributes and can be captured through the LSTM model to capture long-term dependencies in the sequence. However, the LSTM model also has some shortcomings. LSTM is unidirectional and can only process sequential data from the past to the future, and cannot comprehensively analyze temporal data [25, 26].

To comprehensively consider the operational data and meteorological data of the power energy system, a bidirectional LSTM model is utilized, while taking into account both forward and backward information. The main reason for choosing a bidirectional LSTM model is its unique advantage in processing time series data. Bidirectional LSTM can simultaneously capture the forward and backward temporal dependencies of sequence data, which is crucial for comprehensively analyzing the dynamic and complex characteristics of power energy systems. Bidirectional LSTM further enhances the attention to key features through the attention mechanism, and can dynamically adjust the weight of important information in the input sequence, thereby significantly improving classification accuracy and stability.

The structure of the bidirectional LSTM model is illustrated in Fig. 2.

From Fig. 2, the structure of the bidirectional LSTM model is described. The input layer receives temporal data. The bidirectional LSTM model has hidden states in both forward and backward directions. When processing sequence data, not only past information is considered, but future information can also be considered simultaneously, thereby better capturing patterns in the sequence. Connected to the output sequence of bidirectional LSTM through an attention mechanism, forward and backward outputs are operated to capture key information.

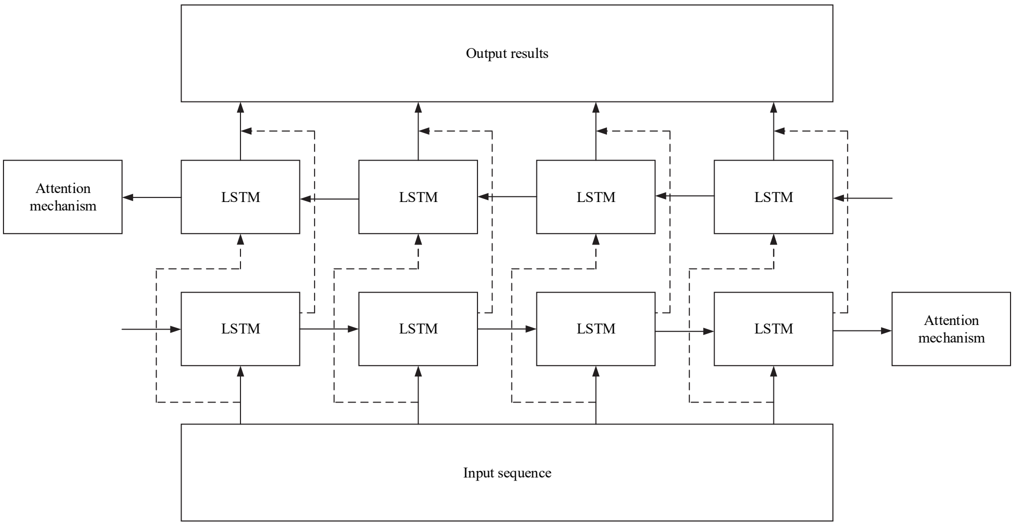


**Figure 2** Structure of bidirectional LSTM model

Attention mechanisms allow models to focus more on different parts of the input when dealing with sequential or aggregate data to better capture relevant information [27, 28]. Through the application of the attention mechanism, the output of each time step depends not only on the input of the current time step but also on different parts of the input sequence. The bidirectional LSTM model can selectively focus on the parts of the input sequence that contribute more to the current output.

The energy at each position t is calculated:

$$e_t = w_t \cdot h_t \tag{6}$$

In Eq. (6), $h_t$ represents the hidden state corresponding to input position $t$, and $w_t$ represents the weight. The energy value is the basis for generating attention weights. By assigning different weights to hidden states, the model can focus on important parts of the input sequence.

Energy $e_t$ is converted into attention weights, resulting in:

$$a_t = \frac{e^{e_t}}{\sum_{t'} e^{e_{t'}}} \tag{7}$$

In Eq. (7), $a_t$ represents attention weight. The function of attention weights is to normalize and ensure that the sum of attention weights for all positions is 1, allowing the model to dynamically allocate the focus of attention. The weighted sum is calculated, and the hidden states of the input sequence are added to the attention weights to obtain:

$$c = \sum_t a_t \cdot h_t \tag{8}$$

Eq. (8), $c$ represents the context vector, and $t$ represents the hidden state at the $t$-th position in the input sequence. The function of weighted summation is that the context vector will contain information about important positions in the input sequence, and the model can generate more meaningful outputs based on this information. The discrepancy between the output results of the bidirectional LSTM model and the actual power energy system leakage is measured by the cross-entropy loss function, which is expressed by the formula:

$$CCL = -\sum_i y_i \cdot \log\left(p_i\right) \tag{9}$$

In Eq. (9), $y_i$ represents the true label of category $i$ in the actual target, $p_i$ represents the probability distribution of category $i$ predicted by the model, and $CCL$ represents the cross entropy loss. By using cross entropy loss, the model can optimize the predicted results to make them closer to the actual target. The training process of the bidirectional LSTM model is to continuously approach the actual vulnerability results of the power energy system until the loss value is minimized. Through the backpropagation algorithm, the gradient of the loss function for the model parameters is calculated, and through the stochastic gradient descent optimization algorithm, the gradient of the current sample is calculated to update the model parameters [29-30].

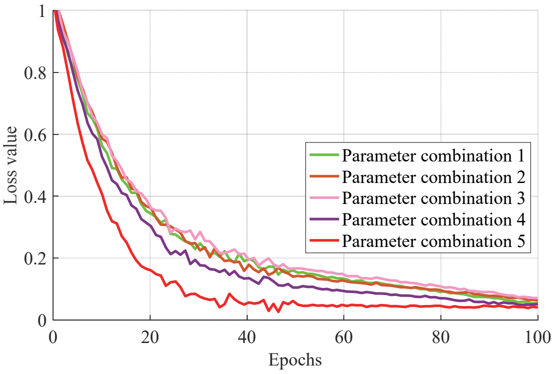


**Figure 3** Loss functions corresponding to different parameter combinations

The loss functions corresponding to different parameter combinations during the training of the bidirectional LSTM model are illustrated in Fig. 3.

In Fig. 3, the loss functions corresponding to different parameter combinations are described. During the training process of the bidirectional LSTM model, the epoch is set to 100. It can be clearly learned that parameter combination 5 represents the smallest loss function value and has a faster convergence speed.

The process of stochastic gradient descent is illustrated in Fig. 4.

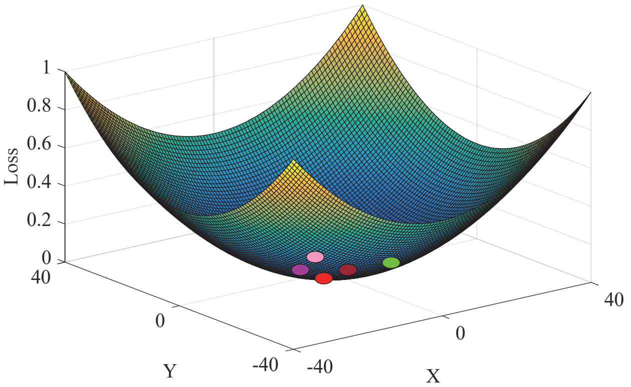


**Figure 4** The process of stochastic gradient descent

In Fig. 4, the process of stochastic gradient descent is described. The *X*-axis and *Y*-axis both show model parameters, while the vertical axis represents loss values. Among them, the scattered points with colors correspond to the results of different parameter combinations. It can be seen that the red scatter is at the minimum point of the entire loss function, and the loss value reaches its lowest point. During the model training process, hyperparameter adjustment is a key step in improving model performance. This study used grid search and experimental comparison methods to optimize hyperparameters, mainly focusing on optimizing learning rate, batch size, hidden units, attention heads, and iteration times (epochs). Firstly, set the candidate parameter range and find the preliminary parameter combination through grid search. Then, based on the preliminary results, further fine-tune key parameters such as learning rate and number of hidden units to ensure that the performance of the model reaches a balance between the training and validation sets.

During the optimization process, the model performance is evaluated through the cross-entropy loss function, and the parameter combination with the minimum loss on the validation set is selected as the final setting. In addition, by observing the learning curve during the training process, it is ensured that the training process is stable and avoids overfitting or underfitting. The final parameter settings are shown in Tab. 4.

**Table 4** Final parameter setting table

| Hyperparameter | Value | TuningDescription |
|---|---|---|
| Learning Rate | 0.003 | Tested in the range of 0.001-0.01 to balance convergence speed and performance. |
| BatchSize | 32 | Selected after testing 16, 32, and 64 to ensure computational efficiency and stability. |
| HiddenUnits | 128 | Chosen after testing 64, 128, and 256 to keep model complexity moderate. |
| Attention Heads | 4 | Tuned based on the feature extraction capability of the attention mechanism. |
| Epochs | 100 | The training curve observed, stabilized after 100 iterations. |
| Optimizer | Adam | Adam optimizer was chosen to accelerate gradient descent. |

### 3.4 Vulnerabilities and Security Protection Strategies

The power energy system is one of the crucial infrastructures in modern society, supporting various production and living activities. As technology advances, power systems are becoming more and more reliant on automation, digitization, and internet technologies, which improve the efficiency of the system, but also introduce new risks and vulnerabilities.

Wind energy, solar energy, and other renewable energy sources are closely related to weather conditions, and extreme weather conditions seriously affect the operation of power system equipment. Comprehensive analysis of meteorological data is beneficial for identifying vulnerabilities in the power energy system.

A bidirectional LSTM model is used to comprehensively analyze meteorological data and operational data of the power energy system, identify potential vulnerabilities in the power energy system, and develop targeted security protection strategies to ensure the security of the power energy system.

The interface for detecting vulnerabilities in the power energy system is shown in Fig. 5.

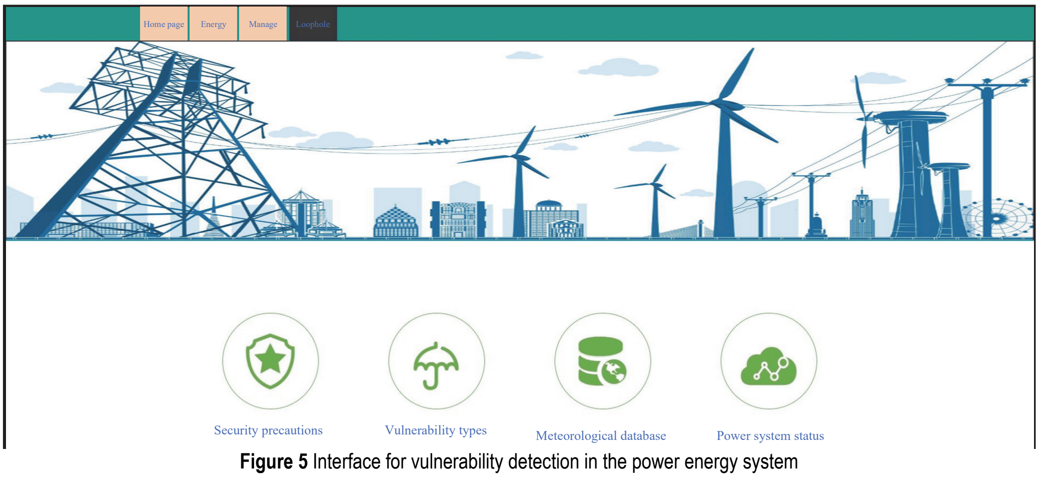


**Figure 5** Interface for vulnerability detection in the power energy system

In Fig. 5, the interface for detecting vulnerabilities in the power energy system is described. The types of vulnerabilities in the power energy system are analyzed using a bidirectional LSTM model, providing timely alerts based on the types of vulnerabilities, and developing security prevention strategies.

The status and performance of power equipment are monitored in real-time. Based on meteorological data, the operating conditions of the equipment under different meteorological conditions are predicted, and preventive maintenance measures are taken. Advanced network security measures are adopted to protect the power system from network attacks and prevent data tampering and leakage.

Security training is provided for power system operators, maintenance personnel, and related management personnel to enhance their understanding of meteorological data vulnerabilities and system security and enhance their ability to respond to emergencies. A detailed emergency response plan is developed, including response strategies for extreme weather events, to ensure that system administrators and operators can respond quickly in emergency situations.

## 4 Performance of Vulnerability Detection and Evaluation in Power Energy Systems

### 4.1 Experimental Environment

Meteorological data has become an indispensable factor in the operation and planning of power systems. The uncertainty and volatility of meteorological conditions pose a series of potential vulnerabilities and security risks to the power energy system.

The instability of the power system caused by increased energy volatility, equipment failures caused by extreme weather events, and power outages caused by insufficient energy supply may all lead to system vulnerabilities. By comprehensively analyzing meteorological data, potential vulnerabilities in the power energy system can be fully analyzed. In order to conduct effective simulation experiments, high-performance computers are used as experimental equipment, with parameters including 32GB of running memory and GeForce RTX 3090.

In Linux, a bidirectional LSTM model is constructed in Python by introducing the TensorFlow deep learning framework. The loss function is set to a cross-entropy loss function, with a model learning rate of 0.003, a batch size of 32, and epochs of 100.

The collected data is divided into datasets to evaluate the performance of vulnerability analysis of power energy systems. The dataset is sourced from real-world power energy systems and meteorological monitoring data. The operational data of the power system is collected through a sensor network of a certain State Grid Corporation of China, including key indicators such as power generation, voltage, current, frequency, and load. The data covers a three-year time span from 2020 to 2022, with records made every hour, totaling over 250000 time series records. Meteorological monitoring data is provided by the National Meteorological Administration and local meteorological stations, covering important meteorological variables such as temperature, humidity, wind speed, and air pressure. This data covers the same time range and is precisely matched with power system operation data through timestamps, forming a multivariate time series dataset. The distribution of the dataset after division is presented in Tab. 5.

**Table 5** Distribution of dataset after partition

| Type | Training set | Testing set |
|---|---|---|
| Physical security vulnerabilities | 3330 | 1110 |
| Network security vulnerabilities | 3420 | 1140 |
| Software security vulnerabilities | 3150 | 1050 |
| Data security vulnerabilities | 3300 | 1100 |
| Internet of Things device vulnerabilities | 2850 | 950 |
| Remote access vulnerability | 3000 | 1000 |
| Normal | 1950 | 650 |

In Tab. 5, the distribution of the dataset after partitioning is described. The original dataset is divided into training and testing sets in a 3:1 ratio. The data in the training set is used for model training, while the data in the testing set is used to evaluate the performance of vulnerability analysis in the power energy system.

### 4.2 Evaluation Indicators

The uncertainty and variability of meteorological conditions have introduced new challenges and vulnerabilities to the power energy system. Meteorological data not only affects energy production but may also lead to instability and vulnerabilities in the power system under extreme weather events. Under strong wind or storm conditions, wind turbines may be damaged or shut down, and solar photovoltaic power systems may experience reduced production capacity under cloud cover. These situations can lead to supply-demand imbalance, voltage instability, or even equipment damage in the power system.

Analyzing the impact of meteorological data on the power system can identify potential vulnerabilities and security risks, help formulate security protection strategies, and enhance the robustness of the power system. The bidirectional LSTM model is utilized to classify vulnerabilities in the power energy system and provide targeted security measures based on the type of vulnerability.

The performance of vulnerability classification in the power energy system is essentially the difference between the predicted vulnerabilities of the bidirectional LSTM model and the actual vulnerability labels. The accuracy, precision, and recall metrics can be calculated based on the predicted vulnerability types and actual vulnerability types in the testing set results.

The accuracy of vulnerability classification is expressed by the formula:

$$Accuracy = \frac{TP + TN}{N} \tag{10}$$

The precision is expressed by the formula:

$$Precision = \frac{TP}{TP + FP} \tag{11}$$

The recall rate is expressed as:

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

To provide a more detailed description of the performance of the bidirectional LSTM model in vulnerability analysis of power energy systems, a confusion matrix is used to visualize the classification findings of specific vulnerability types using the bidirectional LSTM model. To effectively analyze the stability of the bidirectional LSTM model for vulnerability analysis, different categories of vulnerability test samples in the testing set are mixed with normal test samples. By changing the content of the test samples, the changes in vulnerability classification performance are observed.

To ensure the credibility of the research results, the dataset was divided into training sets, validation sets, and testing sets for credibility verification, with proportions of 70%, 15%, and 15%. Adopting 5-fold cross-validation to ensure the robustness of the performance evaluation of the model. Multiple classic models are compared in the experiment, including the bidirectional LSTM model and unidirectional LSTM model, SVM, DT, and RF. The performance of different models for vulnerability classification is visually compared through receiver operating characteristic (ROC) curves.

## 5 RESULTS

### 5.1 Vulnerability Classification Performance

The bidirectional LSTM model was utilized to classify vulnerabilities in power energy systems. The performance of the bidirectional LSTM model for vulnerability classification is illustrated in Fig. 6.

In Fig. 6, the performance of the bidirectional LSTM model for vulnerability classification is described. The horizontal axis represents the data in the testing set being divided into 10 groups, while the vertical axis represents the performance of vulnerability classification. The bidirectional LSTM model can capture contextual information at each time step and automatically extract abstract-level features. The introduction of the attention mechanism makes bidirectional LSTM models more focused on key time steps or features when processing temporal data, thereby improving the model's attention to important information. The bidirectional LSTM model performs well in vulnerability classification of power energy systems, with an average classification accuracy of 97.3%, an accuracy of 97.2%, and a recall rate of 97.8%. The model captures bidirectional temporal dependencies and combines attention mechanisms to focus on key time points and features, greatly improving classification performance. The high accuracy and stability of this model indicate its suitability for complex time-series data analysis, and it has significant advantages compared to traditional methods. This method effectively combines meteorological data and power system operation data, making up for the shortcomings of traditional analysis that ignores meteorological factors. It provides important technical support for real-time monitoring, predictive maintenance, and the development of targeted safety strategies, which helps to improve the reliability and safety of power systems under dynamic meteorological conditions.

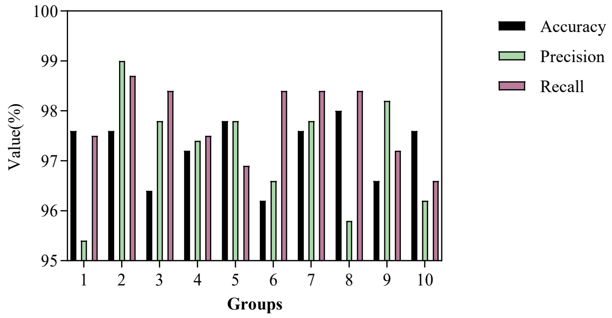
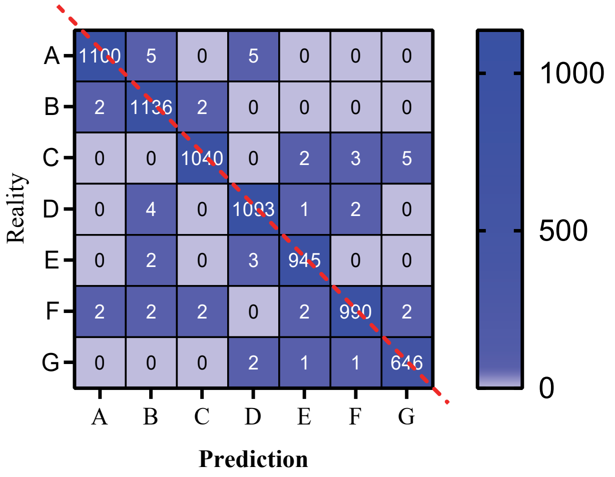


**Figure 6** Vulnerability classification performance

### 5.2 Confusion Matrix

Through the bidirectional LSTM model, vulnerability classification was carried out based on meteorological data and other information. Targeted security measures can be taken according to the types of potential vulnerabilities. The confusion matrix was drawn based on the classification of specific vulnerability types. Physical security vulnerabilities, network security vulnerabilities, software security vulnerabilities, data security vulnerabilities, IoT device vulnerabilities, remote access vulnerabilities, and normal vulnerabilities were respectively recorded as A, B, C, D, E, F, and G. The confusion matrix is shown in Fig. 7.



| | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| A | 1100 | 5 | 0 | 5 | 0 | 0 | 0 |
| B | 2 | 1136 | 2 | 0 | 0 | 0 | 0 |
| C | 0 | 0 | 1040 | 0 | 2 | 3 | 5 |
| D | 0 | 4 | 0 | 1003 | 1 | 2 | 0 |
| E | 0 | 2 | 0 | 3 | 945 | 0 | 0 |
| F | 2 | 2 | 2 | 0 | 2 | 990 | 2 |
| G | 0 | 0 | 0 | 2 | 1 | 1 | 646 |

**Figure 7** Confusion matrix

In Fig. 7, the confusion matrix illustrates the performance of the bidirectional LSTM model in vulnerability classification. Each row represents the actual category of vulnerabilities, while each column represents the predicted category. The red dashed diagonal highlights correctly classified samples, indicating alignment between the predicted and actual ategories. For example, among physical security vulnerabilities, 1100 samples were correctly classified, and for normal samples, 646 were accurately identified. This high classification accuracy across multiple categories reflects the model's ability to capture complex relationships in the data. The bidirectional LSTM model, enhanced with attention mechanisms, ensures that critical features within sequential data are prioritized, enabling precise differentiation between vulnerability types. This result demonstrates the robustness of the model in eal-world scenarios, where accurate detection of vulnerabilities is essential for system security. The confusion matrix also validates the practical utility of this approach.

Accurate classification of vulnerabilities, such as physical security or network vulnerabilities, allows targeted preventive strategies to be developed. For instance, recognizing physical security vulnerabilities enables timely reinforcement of infrastructure, while identifying network vulnerabilities facilitates improved cybersecurity measures. Overall, these results highlight the significance of using advanced models like bidirectional LSTM for comprehensive and reliable vulnerability analysis in power energy systems, effectively bridging the gap between data-driven methods and practical security needs.

### 5.3 Vulnerability Classification Stability

In the actual process of vulnerability analysis in power energy systems, there are often significant differences in the types of vulnerabilities. To analyze the stability of the bidirectional LSTM model for vulnerability classification, the types of vulnerabilities in the testing set were combined with normal samples to construct new test content. The stability results of vulnerability classification are shown in Fig. 8.

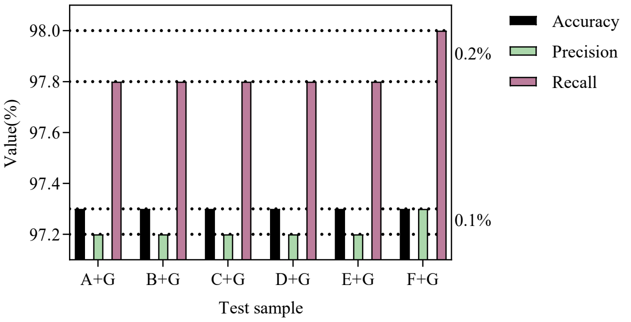


**Figure 8** Vulnerability classification stability

In Fig. 8, the stability of the bidirectional LSTM model in vulnerability classification is evaluated. The horizontal axis represents various combinations of vulnerability test samples with normal samples, while the vertical axis shows classification performance. The results indicate that the model's accuracy remains consistent at 97.3%, with a minimal fluctuation of 0.1%. Similarly, the recall rate ranges from 97.8% to 98.0%, demonstrating a fluctuation of only 0.2%. These results highlight the model's robustness across different testing scenarios. Such stability is critical in real-world applications, where power energy systems face diverse and unpredictable conditions. The minimal performance variation underscores the effectiveness of the bidirectional LSTM model in maintaining reliable classification, regardless of sample composition. This consistency ensures that vulnerability detection remains accurate and dependable, supporting the development of proactive and targeted security measures. Moreover, the high stability further validates the integration of bidirectional temporal dependencies and attention mechanisms, making this approach a strong foundation for enhancing power system resilience under dynamic operational and environmental conditions.

### 5.4 Performance of Different Models

To fully analyze the performance of bidirectional LSTM models for vulnerability classification in power energy systems, multiple classic models were set up for comparison. The performance of different models for vulnerability classification is displayed in Tab. 6.

**Table 6** Vulnerability classification performance of different models

| Model | Accuracy / % | Precision / % | Recall / % | F1 value / % |
|---|---|---|---|---|
| Bidirectional LSTM | 97.3 | 97.2 | 97.8 | 97.5 |
| Unidirectional LSTM | 93.2 | 92.8 | 91.8 | 92.3 |
| SVM | 87.6 | 84.5 | 83.2 | 83.8 |
| DT | 82.1 | 70.8 | 76.3 | 73.4 |
| RF | 84.1 | 78.9 | 74.8 | 76.8 |
| Clustering - State Prediction | 94.5 | 96.1 | 91.8 | 92.0 |
| Adaptive Unscented Kalman Filter Dynamic | 96.1 | 95.2 | 94.4 | 94.8 |
| LSTM | 90.1 | 93.0 | 88.6 | 84.9 |

As shown in Tab. 6, the bidirectional LSTM model achieves the highest performance among all models in vulnerability classification for power energy systems, with an accuracy of 97.3%, precision of 97.2%, recall of 97.8%, and F1-score of 97.5%. In contrast, traditional models like unidirectional LSTM (accuracy 93.2%) and SVM (accuracy 87.6%) exhibit lower performance, highlighting the superiority of the bidirectional approach. The bidirectional LSTM's ability to simultaneously capture forward and backward temporal dependencies provides a more comprehensive understanding of sequential data, crucial for identifying vulnerabilities influenced by dynamic operational and environmental factors. Additionally, the attention mechanism enhances this model's performance by focusing on key time steps, improving its ability to detect subtle patterns that signify vulnerabilities. These results demonstrate that integrating bidirectional temporal analysis with attention mechanisms significantly improves the model's ability to process complex, multivariate time-series data. This advancement ensures more accurate and reliable vulnerability detection, supporting the development of targeted preventive strategies and enhancing the resilience of power energy systems under varying conditions. The consistent outperformance of the bidirectional LSTM model underscores its importance as a tool for robust vulnerability analysis in critical infrastructure systems.

### 5.5 ROC Curve

According to different models, the results of vulnerability classification in the power energy system were carried out. The comparison of the ROC curves drawn is shown in Fig. 9.

Based on Fig. 9, the comparison of ROC curves for various models illustrates their classification capabilities for vulnerabilities in the power energy system. The diagonal line connecting the bottom-left and top-right corners acts as a baseline, representing a random classification scenario where positive and negative examples cannot be effectively distinguished. The ROC curves of all five models are positioned above this baseline, confirming their ability to classify vulnerabilities to some degree. Among these models, the bidirectional LSTM model exhibits the best performance, as its ROC curve is closest to the top-left corner, with the largest area under the curve (AUC). This indicates its superior ability to discriminate between different types of vulnerabilities in the power energy system. The significance of this metric lies in the robustness and accuracy it conveys about the bidirectional LSTM model. The larger AUC reflects its advanced capability in integrating operational and meteorological data, effectively capturing bidirectional temporal dependencies, and leveraging the attention mechanism to focus on critical features. From a research perspective, this strong performance underscores the practicality of combining deep learning techniques with domain-specific knowledge to address real-world challenges in infrastructure vulnerability analysis. The bidirectional LSTM model's ability to provide precise and reliable classification results contributes significantly to improving the safety and stability of power systems. By identifying vulnerabilities with high accuracy, it enables the development of targeted mitigation strategies, reducing risks from extreme weather and other operational uncertainties. This approach not only enhances the resilience of power systems but also demonstrates the potential of advanced machine learning frameworks in critical infrastructure protection.

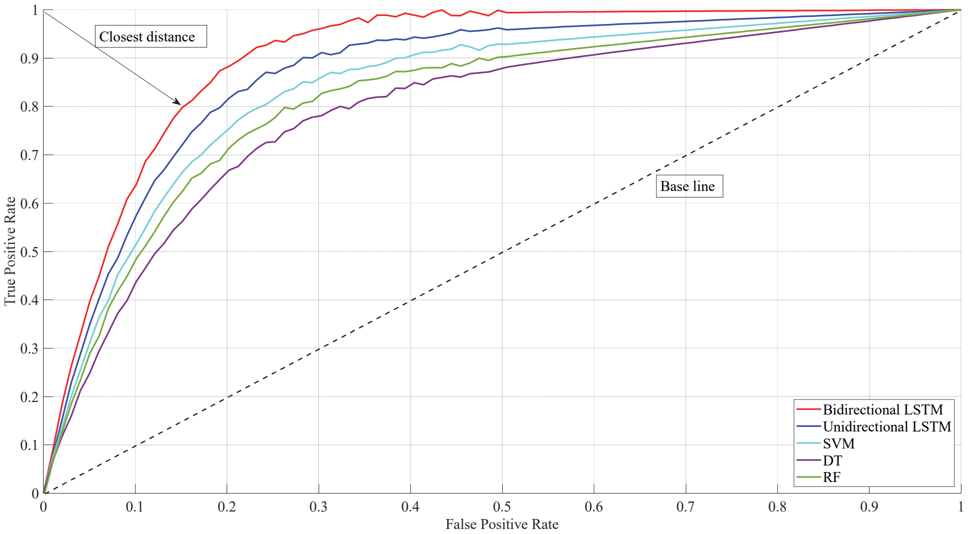


**Figure 9** Comparison of ROC curves

## 6 CONCLUSIONS

The power energy system is one of the key infrastructures in modern society. Meteorological factors such as temperature, humidity, and wind speed can affect the performance of various equipment in the power system, leading to potential safety issues. To effectively analyze vulnerabilities in the power energy system, a bidirectional LSTM model was utilized to comprehensively analyze operational and meteorological data of the power energy system and to consider the output of each time step through attention mechanism weighting. The results indicated that the bidirectional LSTM model could accurately classify vulnerabilities in different types of power energy systems and had high stability in different testing environments. Targeted security protection strategies based on the types of vulnerabilities in the power energy system can effectively ensure the safe and stable operation of the power energy system. The research provides reliable technical support for real-time monitoring and early warning in dynamic environments. This method can be applied to the actual operation of the power system, by accurately identifying vulnerability types and developing targeted security protection strategies, thereby enhancing the system's resilience and security, and responding to complex and changing weather and operating conditions. However, the study did not specifically consider the special factors of extreme weather during design and testing. In the future, optimization of extreme weather and special meteorological conditions will be carried out to expand the applicability of research methods. Moreover, the research methods currently only consider the content and operation of the technology itself, without considering the operational limitations and compatibility issues in complex power systems. In the future, specific and complete power system software and hardware equipment will be considered to optimize the implantation and operation methods of research methods, to more effectively apply them in practical environments and contribute to the greater resilience and security of the power system.

## 7 REFERENCES

[1] Pang, J. L. (2023). Adaptive Fault Prediction and Maintenance in Production Lines using Deep Learning. *International Journal of Simulation Modelling (IJSIMM)*, *22*(4). https://doi.org/10.2507/IJSIMM22-4-CO20

[2] Zhang, L. J., Yang, S. J., Wang, S. J., Zeng, Y. M., Hua, W. C., & Li, G. L. (2023). High - Speed Bearing Dynamics and Applications in Production Lines. *International Journal of Simulation Modelling (IJSIMM)*, *22*(4).

https://doi.org/10.2507/IJSIMM22-4-CO17

[3] Talukder, M. A., Sharmin, S., Uddin, M. A., Islam, M. M., & Aryal, S. (2024). MLSTL - WSN: Machine Learning - Based Intrusion Detection Using SMOTETomek in WSNs. *International Journal of Information Security*, *23*(3), 2139-2158. https://doi.org/10.1007/s10207-024-00833-z

[4] Kim, K., Lee, J. H., Lim, H. K., Oh, S. W., & Han, Y. H. (2022). Deep RNN - Based Network Traffic Classification Scheme in Edge Computing System. *Computer Science and Information Systems*, *19*(1), 165-184. https://doi.org/10.2298/CSIS200424038K

[5] Sahoo, B. & Samantaray, S. R. (2020). System integrity protection scheme for enhancing backup protection of transmission lines. *IEEE Systems Journal*, *15*(3), 4578-4588. https://doi.org/10.1109/JSYST.2020.3013896

[6] Bedi, G., Venayagamoorthy, G. K., Singh, R., Brooks, R. R., & Wang, K. C. (2018). Review of Internet of Things (IoT) in Electric Power and Energy Systems. *IEEE Internet of Things Journal*, *5*(2), 847-870. https://doi.org/10.1109/JIOT.2018.2802704

[7] Venkatanagaraju, K., Biswal, M., & Malik, O. P. (2020). Adaptive third - zone distance protection scheme for power system critical conditions. *IEEE Transactions on Power Delivery*, *36*(3), 1401-1410. https://doi.org/10.1109/TPWRD.2020.3008418

[8] Zemin, L. (2023). Improved Electricity Portfolio Prediction Based on Optimized Ant Colony Algorithm. *Tehnički vjesnik*, *30*(2), 458-464. https://doi.org/10.17559/TV - 20221007130139

[9] Calik Bayazit, E., KoraySahingoz, O., & Dogan, B. (2023). Deep Learning - Based Malware Detection for Android Systems: A Comparative Analysis. *Tehnički vjesnik*, *30*(3), 787-796. https://doi.org/10.17559/TV - 20220907113227

[10] Alhussien, N., Aleroud, A., Melhem, A., & Khamaiseh, S. Y. (2024). Constraining Adversarial Attacks on Network Intrusion Detection Systems: Transferability and Defense Analysis. *IEEE Transactions on Network and Service Management*, *21*(3), 2751-2772. https://doi.org/10.1109/TNSM.2024.3357316

[11] Morovati, B., Lashgari, R., Hajihasani, M., & Shabani, H. (2023). Reduced Deep Convolutional Activation Features (R - DeCAF) in Histopathology Images to Improve the Classification Performance for Breast Cancer Diagnosis. *Journal of Digital Imaging*, *36*(6), 2602-2612. https://doi.org/10.1007/s10278-023-00887-w

[12] Kumar, K. & Pande, B. P. (2023). Air pollution prediction with machine learning: a case study of Indian cities. *International Journal of Environmental Science and Technology*, *20*(5), 5333-5348. https://doi.org/10.1007/s13762-022-04241-5

[13] Poddar, A., Gupta, P., Kumar, N., Shankar, V., & Ojha, C. S. P. (2021). Evaluation of reference evapotranspiration methods and sensitivity analysis of climatic parameters for sub - humid sub - tropical locations in western Himalayas (India). *ISH Journal of Hydraulic Engineering*, *27*(3), 336-346. https://doi.org/10.1080/09715010.2018.1551731

[14] Rahman, O., Muttaqi, K. M., & Sutanto, D. (2019). High Temperature Superconducting Devices and Renewable Energy Resources in Future Power Grids: A Case Study. *IEEE Transactions on Applied Superconductivity*, *29*(2), 1-4. https://doi.org/10.1109/TASC.2019.2895677

[15] Staffell, I., Scamman, D., Abad, A. V., Balcombe, P., Dodds, P. E., Ekins, P. et al. (2019). The Role of Hydrogen and Fuel Cells in the Global Energy System. *Energy & Environmental Science*, *12*(2), 463-491. https://doi.org/10.1039/C8EE01157E

[16] Gür, T. M. (2018). Review of Electrical Energy Storage Technologies, Materials, and Systems: Challenges and Prospects for Large - Scale Grid Storage. *Energy & Environmental Science*, 11(10), 2696-2767. https://doi.org/10.1039/C8EE01419A

[17] Moradi, M., Kordestani, M., Jalali, M., Rezamand, M., Mousavi, M., Chaibakhsh, A., & Saif, M. (2024). Sensor and Decision Fusion - Based Intrusion Detection and Mitigation Approach for Connected Autonomous Vehicles. *IEEE Sensors Journal*, *24*(13), 20908-20919. https://doi.org/10.1109/JSEN.2024.3397966

[18] Zhao, F., Dong, B., Pan, H., & Shi, A. (2023). A mining algorithm to improve LSTM for predicting customer churn in railway freight traffic. *Studies in Informatics and Control*, *32*(2), 25-38. https://doi.org/10.24846/v32i2y202303

[19] Imrana, Y., Xiang, Y., Ali, L., Noor, A., Sarpong, K., & Abdullah, M. A. (2024). CNN - GRU - FF: A Double - Layer Feature Fusion - Based Network Intrusion Detection System Using Convolutional Neural Network and Gated Recurrent Units. *Complex & Intelligent Systems*, *10*(3), 3353-3370. https://doi.org/10.1007/s40747-023-01313-y

[20] Maier, M. I., Czibula, G., & Delean, L. R. (2023). Using Unsupervised Learning for Mining Behavioural Patterns from Data. A Case Study for the Baccalaureate Exam in Romania. *Studies in Informatics and Control*, *32*(2), 73-84. https://doi.org/10.24846/v32i2y202307

[21] Kinaci, B. F. & Özarpa, C. (2023). Safety Calculation Model of Grade Crossings with Automatic Barrier System. *Tehnički vjesnik*, *30*(2), 642-647. https://doi.org/10.17559/TV - 20220613220928

[22] Hassler, S. C., Mughal, U. A., & Ismail, M. (2024). Cyber - Physical Intrusion Detection System for Unmanned Aerial Vehicles. *IEEE Transactions on Intelligent Transportation Systems*, *25*(6), 6106-6117. https://doi.org/10.1109/TITS.2023.3339728

[23] Shewalkar, A., Nyavanandi, D., & Ludwig, S. A. (2019). Performance Evaluation of Deep Neural Networks Applied to Speech Recognition: RNN, LSTM, and GRU. *Journal of Artificial Intelligence and Soft Computing Research*, *9*(4), 235-245. https://doi.org/10.2478/jaiscr-2019-0006

[24] Kunang, Y. N., Nurmaini, S., Stiawan, D., &Suprapto, B. Y. (2024). An End - to - End Intrusion Detection System with IoT Dataset using Deep Learning with Unsupervised Feature Extraction. *International Journal of Information Security*, *23*(3), 1619-1648. https://doi.org/10.1007/s10207-023- 00807-7

[25] Sahoo, B. B., Jha, R., Singh, A., & Kumar, D. (2019). Long Short - Term Memory (LSTM) Recurrent Neural Network for Low - Flow Hydrological Time Series Forecasting. *Acta Geophysica*, *67*(5), 1471-1481. https://doi.org/10.1007/s11600-019-00330-1

[26] Lamurias, A., Sousa, D., Clarke, L. A., & Couto, F. M. (2019). BO - LSTM: Classifying Relations via Long Short - Term Memory Networks along Biomedical Ontologies. *BMC Bioinformatics*, *20*(1), 1-12. https://doi.org/10.1186/s12859-018-2584-5

[27] Eljialy, A. E. M., Uddin, M. Y., & Ahmad, S. (2024). Novel Framework for an Intrusion Detection System Using Multiple Feature Selection Methods Based on Deep Learning. *Tsinghua Science and Technology*, *29*(4), 948-958. https://doi.org/10.26599/TST.2023.9010032

[28] Singh, I. & Jindal, R. (2024). Outlier Based Intrusion Detection in Databases for User Behaviour Analysis Using Weighted Sequential Pattern Mining. *International Journal of Machine Learning and Cybernetics*, *15*(7), 2573-2593. https://doi.org/10.1007/s13042-023-02049-4

[29] Netrapalli, P. (2019). Stochastic Gradient Descent and Its Variants in Machine Learning. *Journal of the Indian Institute of Science*, *99*(2), 201-213. https://doi.org/10.1007/s41745-019-0098-4

[30] Jotić, G., Štrbac, B., Toth, T., Blanuša, V., Dovica, M., & Hadžistević, M. (2023). The Analysis of Metrological Characteristics of Different Coordinate Measuring Systems. *Tehnički vjesnik*, *30*(1), 32-38. https://doi.org/10.17559/TV-20220204091212

**Contact information:**

**Hongxia WANG**
State Grid Xinjiang Company Limited Electric Power Research Institute,
Urumqi 830000, Xinjiang, China
Xinjiang Key Laboratory of Extreme Environment Operation and Testing Technology for Power Transmission & Transformation Equipment,
Urumqi 830000, Xinjiang, China
E-mail: wanghongxiaab@163.com

**Jie XU**
(Corresponding author**)**
Xinjiang Meteorological Service Center,
Urumqi 830002, Xinjiang, China
E-mail: basifei007170@sohu.com

**Yang YANG**
State Grid Xinjiang Company Limited Electric Power Research Institute,
Urumqi 830000, Xinjiang, China
Xinjiang Key Laboratory of Extreme Environment Operation and Testing Technology for Power Transmission & Transformation Equipment,
Urumqi 830000, Xinjiang, China

**Meng LI**
State Grid Xinjiang Company Limited Electric Power Research Institute,
Urumqi 830000, Xinjiang, China
Xinjiang Key Laboratory of Extreme Environment Operation and Testing Technology for Power Transmission & Transformation Equipment,
Urumqi 830000, Xinjiang, China