MRMR Feature Selection Algorithm for Microgrid Frequency Stability Classification

Ngoc Au Nguyen
Ho Chi Minh City University of Technology and Education, Vietnam
aunn@hcmute.edu.vn (corresponding author)

Nghia Le Trong
Ho Chi Minh City University of Technology and Education, Vietnam
trongnghia@hcmute.edu.vn

Duy Bui Cong
Ho Chi Minh City University of Technology and Education, Vietnam
congduy0908@gmail.com

Phuong Nam Nguyen
Cao Thang Technical College, Vietnam | Ho Chi Minh City University of Technology and Education, Vietnam
nguyenphuongnam@caothang.edu.vn

Received: 24 February 2025 | Revised: 2 April 2025 and 7 April 2025 | Accepted: 9 April 2025

 DOI: https://doi.org/10.48084/etasr.10711

ABSTRACT

This paper proposes applying the Minimum Redundancy Maximum Relevance (MRMR) algorithm to select variables for constructing a deep neural network-based classifier for Microgrid (MG) frequency stability assessment. The MRMR algorithm is combined with the 1-Nearest Neighbor (1-NN) machine learning classifier (MRMR&1-NN) to evaluate the classification accuracy in feature selection. The study also compares MRMR-based feature selection with Fisher, Relieff, and Chi-squared methods. Reducing the feature space is crucial for minimizing computational cost, optimizing memory usage, and reducing the expenses of sensor measurement equipment in practical applications. Experimental results on a 16-bus MG system demonstrate that the proposed method not only significantly reduces the number of inputs, but also improves the classification accuracy. The MRMR method achieves higher accuracy compared to the other feature selection techniques. Based on the MRMR&1-NN feature selection results, this paper proposes employing a Bidirectional Long Short-Term Memory network with Fully Connected layers (BiLSTM-FC) for model construction. The results indicate that the BiLSTM-FC model achieves high classification accuracy, highlighting the effectiveness of using MRMR for feature selection and applying the BiLSTM-FC classifier for MG frequency stability classification.

Keywords-feature selection; frenquency stability classification; microgird; neural networks; deep learning

I. INTRODUCTION

Early detection of frequency instability in power systems is crucial for timely control actions to maintain system stability. In the event of a fault, a Microgrid (MG) is highly susceptible to frequency instability. This issue arises because MGs rely heavily on renewable energy sources [1]. Renewable energy sources, such as solar and wind power, offer the advantage of being environmentally friendly; however, they are highly dependent on weather conditions. Consequently, in the event of a failure, weather fluctuations can lead to supply shortages, causing power imbalances that result in voltage and frequency instability [2], potentially leading to MG collapse. Clearly, since MGs are inherently sensitive to variations in renewable energy sources, predicting frequency stability becomes even more essential.

The Artificial Neural Network (ANN) method plays a crucial role in rapidly analyzing data [3] and enabling early prediction of frequency instability in MGs. Traditional analytical methods fail to meet this requirement, making machine learning approaches, such as neural networks, essential for addressing this challenge [4, 5]. Recently, numerous studies have focused on applying deep neural networks. In [6, 7], Convolutional Neural Networks (CNNs) were utilized to assess the voltage stability of power systems.

In [8], deep neural networks were applied for power system stability assessment. Authors in [9] proposed an embedded Long Short-Term Memory (LSTM) model optimized using Bayesian techniques to predict the next-day Photovoltaic (PV) power output. LSTM, a deep learning method, is widely used in the field of machine learning. LSTM is a type of Recurrent Neural Network (RNN) designed to address the challenge of learning and storing long-term dependencies in sequential data. It can capture long-range relationships within time-series data, overcoming the vanishing gradient problem encountered by traditional RNNs. LSTM achieves long-term memory retention through a specialized gating mechanism, which includes the forget gate, input gate, and output gate. This mechanism enables LSTM to retain and adjust important information throughout the training process, thereby improving its ability to predict time series or problems that rely on long-term context. LSTM memorizes long-term information and helps predict future values based on past information. However, LSTMs can still encounter the vanishing gradient problem (although to a lesser extent than traditional RNNs), and in certain cases, information from the beginning of the sequence may not be well preserved during training. Bidirectional Long Short-Term Memory (BiLSTM) is a variant of LSTM designed to enhance learning and prediction of sequential data by utilizing information from both forward and backward directions. BiLSTM can achieve higher prediction accuracy compared to standard LSTM, as it considers both past and future contexts when making predictions.

When applying ANN to power system stability assessment, processing input data attributes or reducing the number of input attributes has practical significance in minimizing the number of measurement sensors required. Intelligent search methods provide high accuracy in selecting the optimal set of input attributes; however, these algorithms are often complex and computationally expensive. Feature space reduction for power system stability assessment models has been proposed by many researchers [10-13]. The Relief method has been proposed for input feature reduction in neural networks for power system security assessment in [10]. In [11], the Sequential Forward Selection (SFS) algorithm was applied to select input variables for power system security classification models. The Fisher method was used in [12] to reduce variables in power system security assessment and identification. Additionally, the Relief, Fisher, and Chi-squared methods have been suggested for selecting input features in intelligent systems for power system stability assessment [13]. In [14], the Binary Particle Swarm Optimization (BPSO) algorithm was applied for input feature selection in power system stability classification models. These studies consistently emphasize the need to reduce the number of input features. Published studies on feature selection for power system stability assessment models indicate that applying statistical ranking-based methods for input selection is computationally simple and easy to implement. In [14], it is shown that the classification accuracy depends on a specific set of variables, the correlation between these variables, and the correlation between input variables and the target output. The study in [15] proposes the Minimum Redundancy Maximum Relevance (MRMR) algorithm to minimize redundancy among input variables while maximizing their relevance to the target output. In [16], mutual information was used for feature selection in speech emotion recognition using machine learning algorithms. Additionally, authors in [17] proposed the application of MRMR algorithm for heart disease diagnosis.

This paper studies the selection of input variables for the neural network model for classifying the frequency stability of MG power grids. Specifically, the paper introduces the application of the MRMR algorithm to reduce the variable space and proposes the use of the deep learning BiLSTM neural network for classifying the frequency stability of MG power grids. The study evaluates its performance on a 16-bus MG system.

II. MINIMUM REDUNDANCY MAXIMUM RELEVANCE ALGORITHM

The MRMR algorithm, introduced in [14], is a feature selection method that aims to select the most relevant features while minimizing redundancy. It ensures that the selected features contain the maximum information about the target variable without having high correlations among themselves.

The objective of the MRMR algorithm is to find an optimal set S of features such that the correlation between S and the output, y, is maximized while the redundancy within S is minimized. The correlation and redundancy are determined using the mutual information, I. I between two variables measures the extent to which the uncertainty of one variable can be reduced by knowing the other. The mutual information between two discrete random variables X and Z is defined as in (1).

$$I(X,Z) = \sum_{i,j} P(\mathrm{X} = \mathrm{x}_i, \mathrm{Z} = \mathrm{z}_j) \cdot \log \frac{P(X = x_i, Z = z_j)}{P(X = x_i)\,\mathrm{P}(Z = z_j)} \quad (1)$$

If X and Z are independent, then $I = 0$. If X and Z are identical random variables, then I equals the entropy of X.

The MRMR algorithm ranks features using a forward selection approach based on the mutual information quotient. This value is referred to as the $MRMR_{score}$, as defined in (2), where $|S|$ represents the number of features in the set S.

$$MRMR_{score} = \frac{I(x,y)}{\frac{1}{|S|}\sum_{z \in S} I(x,\mathrm{z})} \quad (2)$$

The *fscmrmr* function, available in MATLAB 2021a, ranks all features in the set Ω and returns *idx* by the MRMR algorithm. *idx* represents the feature indices ranked in order of significance. This function quantifies feature importance using a heuristic approach and returns a score value, where a higher score indicates greater feature importance. The *fscmrmr* function executes the MRMR algorithm to rank the variables in the following sequence:

- Step 1: Select the variable with the highest relevance, $\max\{[I(x, y)], x \in \Omega\}$. Add the selected feature to the initially empty set S.

- Step 2: Find the variables in the complement set S_c that have nonzero relevance and zero redundancy. If S_c does not contain any variables that satisfy the above conditions, proceed to Step 4. Otherwise, select the variable with the highest relevance, $\max\{[I(x, y)], x \in S_c, W = 0\}$. Add the selected variable to the set S.
- Step 3: Repeat step 2 until the redundancy is non-zero for all variables in S_c.
- Step 4: Select the variable with the largest $MRMR_{score}$ with non-zero relevance and redundancy in S_c, and add the selected variable to the set S.
- Step 5: Repeat step 4 until the relevance is zero for all variables in S_c.
- Step 6: Add the variables with zero relevance to S in a random order.

III. DEEP NEURAL NETWORK BILSTM

A. LSTM Network

The working principle of an LSTM unit is based on a control mechanism through four "gates": the forget gate, the input gate, a tanh layer, and the output gate [9]. These gates help the LSTM determine which information should be retained and which should be forgotten during sequence data processing. The structure of an LSTM unit is illustrated in Figure 1 and the control mechanisms are described as follows.

- The forget gate determines which part of the cell state information C_{t-1} from the previous step should be "forgotten" in the current step. It acts as a filter, allowing only important information to be retained. The Forget Gate is calculated using (3).

$$f_t = \sigma\left(W_f.\left[h_{t-1}, x_t\right] + b_f\right) \quad (3)$$

where f_t is the forget gate value, σ is the sigmoid activation function, W_f represents the weights and bias of the forget gate, h_{t-1} is the output of the LSTM at the previous time step, x_t is the input at the current time step.

- The input gate determines which information will be added to the cell state C_t of the LSTM. The input gate consists of two parts. A sigmoid activation gate decides which portion of the new information should be written into the cell state. A tanh gate generates new candidate values for the cell state. The input gate is calculated using (4) and (5).

$$i_t = \sigma\left(W_i.\left[h_{t-1}, x_t\right] + b_i\right) \quad (4)$$

$$\hat{C}_t = \tanh\left(W_C.\left[h_{t-1}, x_t\right] + b_C\right) \quad (5)$$

where i_t is the input gate value, $\hat{C}_t$ is the candidate values added to the cell state, W_i and b_i are the weights and bias of the input gate, W_c and b_c are the weights and bias of the updated candidate values.

- Cell state update: The cell state C_t serves as the long-term memory storage in an LSTM. It is updated by combining old information (filtered through the forget gate) and new information (introduced via the input gate). The cell state update is computed using (6).

$$C_t = f_t.C_{t-1} + i_t.\hat{C}_t \quad (6)$$

- The output gate determines which information from the cell state C_t will be transmitted as the output. It utilizes a sigmoid activation function to control the extent of information released from the cell state. The cell state is then processed through a tanh function for normalization. The output gate is computed using (7) and (8).

$$o_t = \sigma\left(W_o.\left[h_{t-1}, x_t\right] + b_o\right) \quad (7)$$

$$h_t = o_t.\tanh(c_t) \quad (8)$$

where o_t is the output gate value and h_t is the output of the LSTM at the current time step.

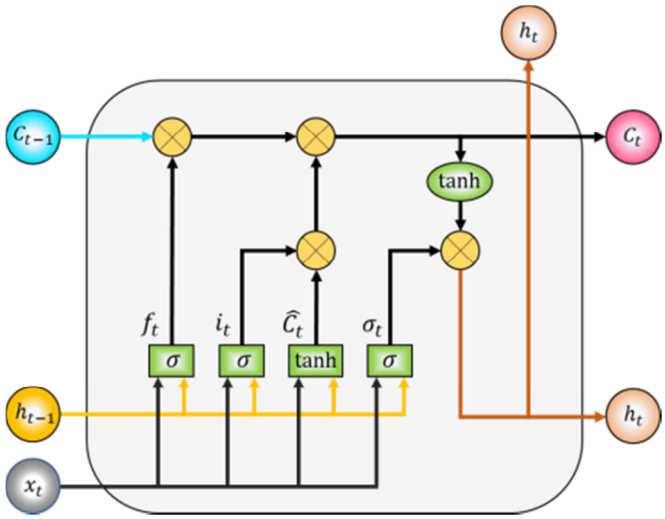

Fig. 1. LSTM network structure.

B. BiLSTM Network

BiLSTM [18] is a variant of LSTM in which sequence data are processed in forward and backward direction. The BiLSTM network enhances the model's ability to understand the context of the sequential sample more effectively. The structure of the BiLSTM network is illustrated in Figure 2, which includes key components such as the input layer, forward processing layer, and backward processing layer. These are specified as follows:

- Input: A sequence of data is provided to the BiLSTM network, denoted as x_T.
- Forward pass processing: The sequential data are passed through an LSTM in the forward direction, from left to right. At each time step t, the forward LSTM processes the input x_t and updates the hidden state h_t along with the cell state C_t.
- Backward pass processing: Simultaneously, BiLSTM employs another LSTM to process the sequence in the reverse direction, from right to left. The backward LSTM updates the hidden state and cell state similarly to the

forward LSTM, but the data are processed in the opposite order.

- Merging information from both directions: Finally, the information from both the forward and backward LSTMs is combined. Typically, the outputs from both directions are concatenated to form the BiLSTM output at time step *t*. The BiLSTM output h_t^{BiLSTM} is represented as $h_t^{BiLSTM} = \left[h_t^{forward}, h_t^{backward}\right]$, where $h_t^{forward}$ is the output from the forward LSTM, $h_t^{backward}$ is the output from the backward LSTM.

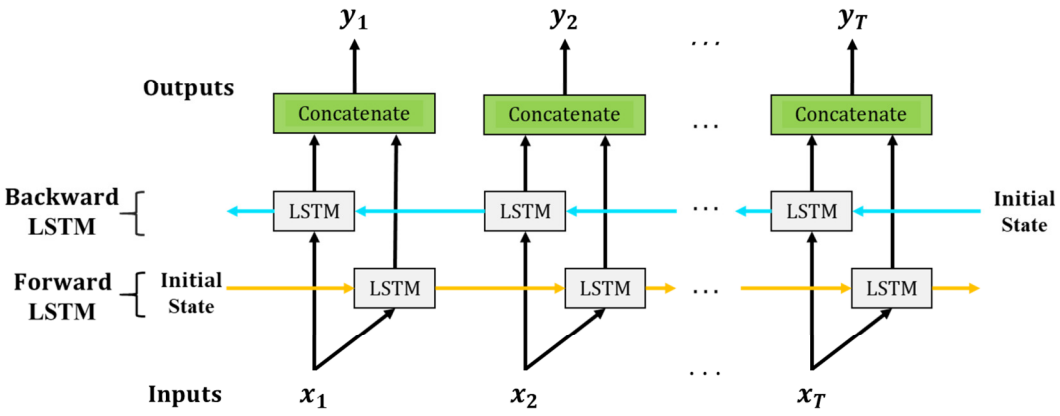

Fig. 2. BiLSTM network structure.

IV. MODEL DESCRIPTION

A. Schematic Selection and Data Collection

To build an ANN model for frequency stability classification of the MG power grid, the first step is to select the schematic and proceed with data collection. The data can be collected through a data acquisition monitoring system or via simulation using software. In this study, the data are generated from simulations with the support of PowerWorld software. The collected data related to the frequency stability of the MG should consider abnormal situations that may occur, such as a generator failure causing a loss of power supply; the MG operating and connected to the large power system, experiencing a fault leading to islanding mode; or the MG operating in islanding mode and suddenly losing a generator. To ensure comprehensive coverage of operational modes, it is important to consider the operating load levels of the MG. The data should contain information closely related to the frequency stability status of the MG. In cases of large fluctuations leading to instability in the MG, the data attribute is related to the variation in active (*P*) and reactive (*Q*) power distributed across the transmission lines, the loads, the generation sources; the voltage and frequency at the buses. The components for building the frequency stability classification model of the MG are shown in Figure 3.

B. Input Data Attribute Selection

The initial collected dataset has *n* variables; after going through the variable reduction stage, the number of remaining variables is *m*, *m*<*n*. Choosing or reducing the number of input data features is useful for minimizing the number of measurement devices or sensors required. Reducing the number of input features helps to improve the classification accuracy and reduce the data storage memory. This paper proposes using the MRMR algorithm, as presented earlier, to select the input features for the frequency stability classification model of the MG.

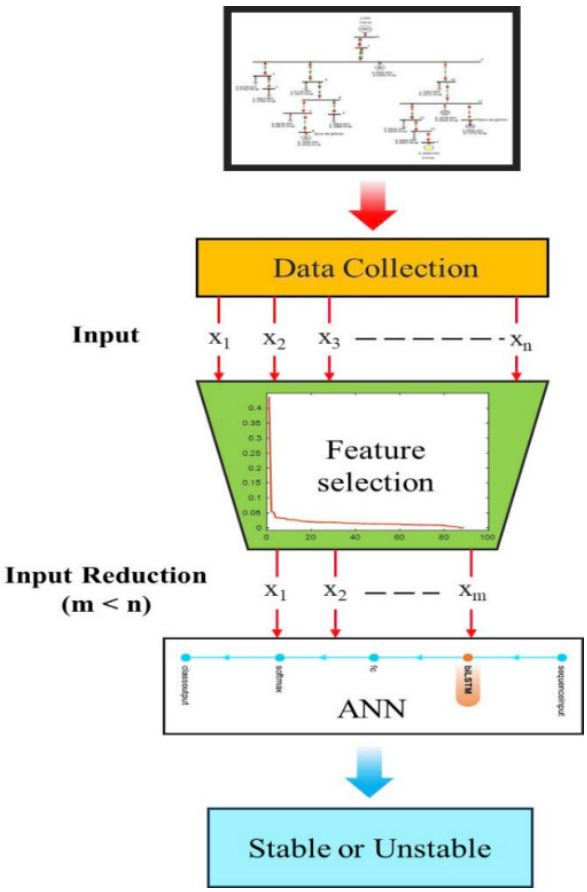

Fig. 3. Components for building the frequency stability classification model of the MG.

C. Building the Neural Network Model

This step involves the construction of an ANN model for frequency stability classification of the MG. MATLAB 2021a provides powerful tools in the Deep Learning Toolbox for building and training BiLSTM networks. This section presents the structure of BiLSTM Fully Connected (BiLSTM-FC) network, which is proposed for constructing the frequency stability classification model of the MG. This configuration will be applied to build the classification model for the 16-bus power grid, as detailed in Section V below.

- sequenceInputLayer: This is the layer that accepts sequential input, suitable for processing sequential data, which is ideal for classification problems. The number of inputs corresponds to the number of input variables in the dataset.
- bilstmLayer: This layer processes the input data in both directions, forward and backward. It performs learning based on BiLSTM networks. The parameter *numHiddenUnits* defines the number of hidden units in each LSTM. The *OutputMode* can be set to '*last*' (to return the output at the last time step only) or '*sequence*' (to return outputs for all time steps). In this case, '*last*' will return the output at the final time step after the data has been processed through both directions.
- fullyConnectedLayer: This layer is used to convert the output of the BiLSTM into a classification layer or target value. The fully connected layer has 2 outputs corresponding to a two-class classification problem, such as the classification of stable or unstable frequency in the MG. In other words, this layer is used to transform the output of the BiLSTM into a binary classification output.
- softmaxLayer: This layer is used for classification tasks. It converts the output of the fully connected layer into probabilities for each class using the softmax function.
- classificationLayer: This is the output layer in a classification network, used to compute the loss function in classification tasks.

D. Evaluation

The training and testing results are evaluated using (9). *CCS* is the sum of correctly classified samples, and *TS* is the sum of samples in the dataset.

$$Acc(\%) = \frac{CCS}{TS}.100 \quad (9)$$

V. APPLYING FREQUENCY STABILITY EVALUATION OF THE MG BASED ON THE BILSTM-FC NETWORK

A. Building the Sample Dataset

In this paper, the data are generated from the simulation of the 16-bus MG system shown in Figure 4. The system consists of 6 generation sources, the generator at bus 16 is considered as the MG connected to the main grid, with 2 diesel generators connected to buses 2 and 8, a solar power source at bus 14, a wind power source at bus 15, and an energy storage battery at bus 11. The model also includes 8 loads located at buses 3, 4, 5, 7, 9, 10, 12, and 13.

To simulate the data collection, two scenarios are considered: the first scenario involves the MG operating in grid-connected mode and a disconnection fault occurs with the main grid; the second scenario is when the MG is operating in islanded mode, considering generator failures sequentially at each of the generation source buses. The simulation process will vary the total load power from 2.1 MW to 6.3 MW, with each load step change being 10%. The load profile is shown in Figure 5. The simulation samples consist of stable samples and unstable samples. The stable samples consist of data collected when the frequency oscillation remains within the allowable range of ±0.3 Hz [19]. Conversely, data that falls outside this range are considered unstable.

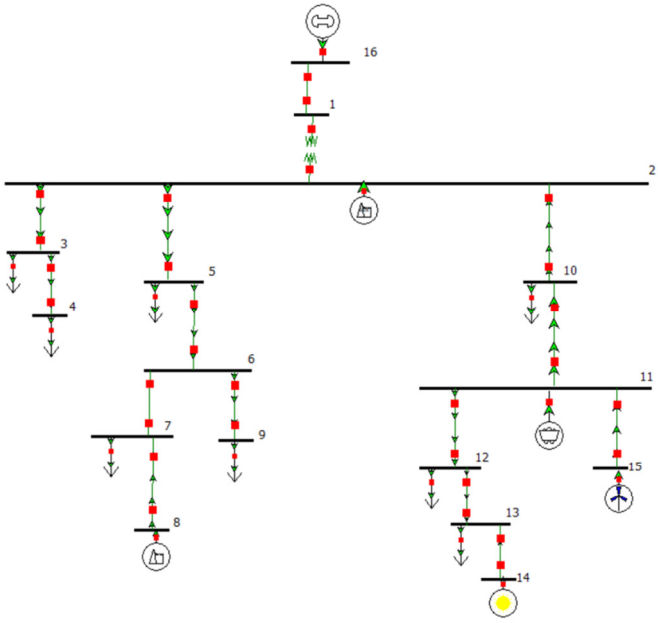

Fig. 4. The 16-bus MG system.

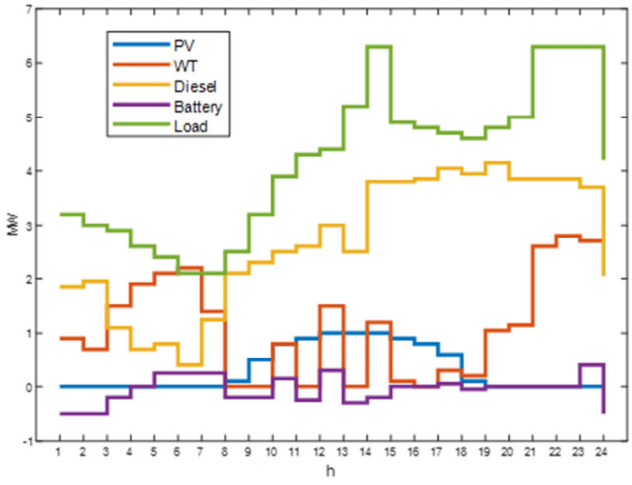

Fig. 5. Load profile.

The data contain information such as the active power deviation and reactive power deviation of the generation source, the active power deviation and reactive power deviation of the load, the frequency deviation at the buses, the voltage deviation at the buses, and the active power deviation and reactive power deviation distributed across the transmission lines. The dataset contains a total of 200 samples, with 110 stable samples and 90 unstable samples. The labels of the stable and unstable samples are 0 and 1. There are 89 input features in the dataset, including: 11 features related to the *P* and *Q* of the generators and the energy storage battery operating in active power generation mode; 32 features related to the frequency and voltage at the buses; 30 features related to the *P* and *Q* distributed across the transmission lines; 16 features related to the *P* and *Q* of the loads.

B. Input Data Feature Processing

This section applies the MRMR algorithm to select input features for the frequency stability classification model of the MG and compares the results with statistical feature selection methods, including Relieff, Chi-squared, and Fisher. The results of calculating the importance scores of the variables and ranking them are presented in Figure 6. The classifier chosen to evaluate the classification accuracy is the 1-Nearest Neighbor (1-NN) classifier. The 1-NN classifier is chosen because it is a simple and easy to implement algorithm. It is constructed using cross-validation with *k-fold* = 10. The dataset is split into 10 equal parts. The model is trained and validated 10 times, each time with a different fold serving as the validation set, whereas the remaining 9 folds are used for training. The results from each of the 10 trials are then averaged to provide a final performance metric. The classification accuracy evaluation results are presented in the graph in Figure 7, where it can be observed that the classification accuracy of 1-NN with the MRMR feature selection method (MRMR&1-NN) is higher than the other methods in the range with up to 20 features. Specifically, with 5 and 15 features, the accuracy reaches 97% and 98%, respectively. This result will be used to proceed with the training of the ANN in the next section. The computations are executed on a laptop with an Intel CORE i7 8565U CPU, 8 GB of memory, and a 1 TB SSD.

C. Network Training

With the number of variables, *m*, equal to 5 and 15, selected using the MRMR method, the paper trains two networks: LSTM-FC and BiLSTM-FC, to compare the training and testing results. The training results of the LSTM-FC and BiLSTM-FC networks are presented in Table I. The structure of the LSTM-FC network is similar to that of the BiLSTM-FC, with the main difference being that the BiLSTM layer is replaced by an LSTM layer.

TABLE I. TRAINING RESULTS FOR LSTM-FC AND BILSTM-FC NETWORKS

Number of inputs	ANN	Training (%)	Testing (%)
5	LSTM-FC	100	87.5
	BiLSTM-FC	100	97.5
15	LSTM-FC	100	95
	BiLSTM-FC	100	97.5

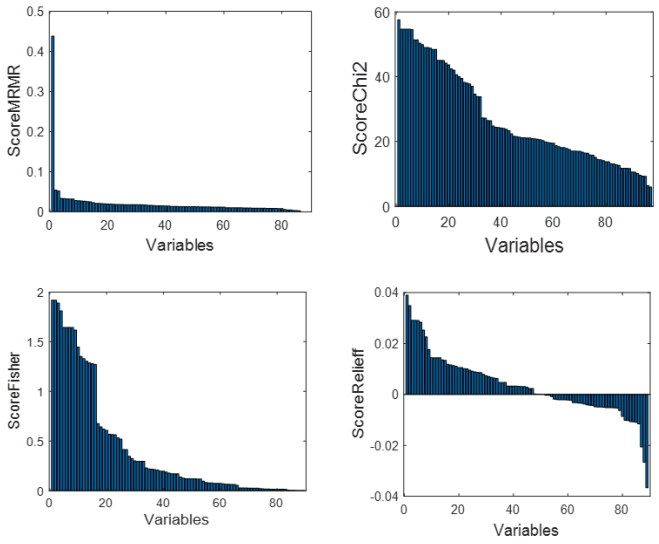

Fig. 6. Chart ranking the importance score of variables.

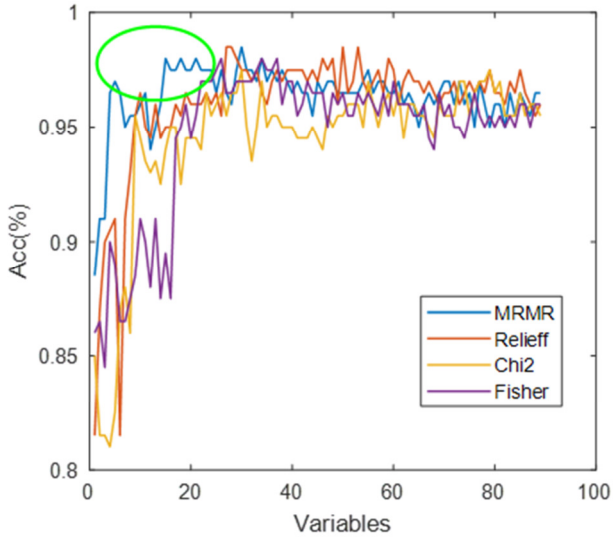

Fig. 7. Feature selection evaluation.

The training accuracy and loss graphs for the LSTM-FC and BiLSTM-FC networks with 5 and 15 variables are presented in Figures 8 and 9. The training is performed using a trial-and-error method. The training algorithm used is '*adam*', with 100 hidden neurons, and the number of iterations for the LSTM-FC network is 100 with *m*=5. For the other cases, the number of iterations is 50. To build the network, the sample set is randomly split into a training set and a testing set. The training set consists of 160 samples, with 88 stable samples and 72 unstable samples. The testing set contains 40 samples, including 22 stable samples and 18 unstable samples.

D. Discussion

According to Figure 7, it is clear that within the range of up to 20 variables, the classification accuracy of 1-NN used to evaluate feature selection shows that the MRMR method outperforms the Relieff, Chi-squared, and Fisher methods. For 5 variables, the feature selection accuracy using MRMR is

97%, whereas the accuracies for the Relieff, Chi-squared, and Fisher methods are 91%, 82.5%, and 89%, respectively. For 15 variables, the feature selection accuracy using MRMR is 98%, whereas the accuracies for the Relieff, Chi-squared, and Fisher methods are 95%, 94%, and 89.5%, respectively.

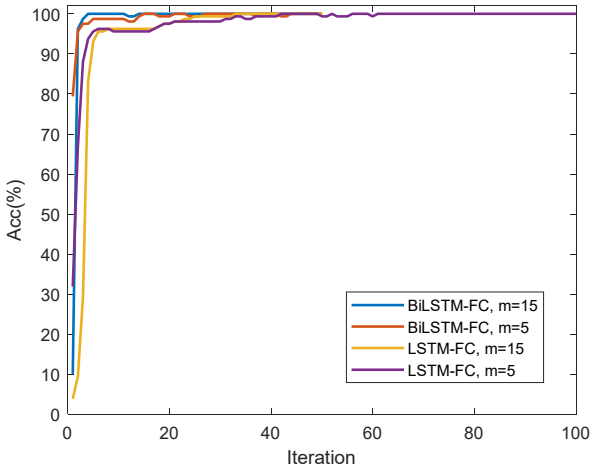

Fig. 8. The training accuracy graph.

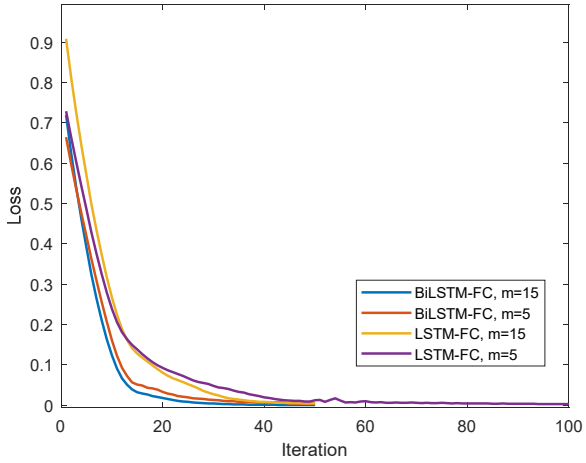

Fig. 9. The training loss graph.

With MRMR&1-NN, the accuracy is 96.5% for 89 variables, whereas the accuracy is 97%for 5 variables and 98% for 15 variables. These results indicate that the MRMR algorithm effectively reduces the number of features. For 5 and 15 variables, the number of features is reduced by 94.4% and 83.4%, respectively, whereas the accuracy increases by 0.5% and 1.5%, respectively.

In Figure 8 and 9, with *m*=5, the training accuracy of the LSTM-FC network reaches 100% around the 42nd iteration and the training loss approaches zero around the 74th iteration; the training accuracy of the BiLSTM-FC network reaches 100% around the 15th iteration, and the training loss approaches zero around the 46th iteration. With *m*=15, the training accuracy of the LSTM-FC network reaches 100% at the 33rd iteration and the training error approaches zero around the 48th iteration; the training accuracy of the BiLSTM-FC network reaches 100% at the 4th iteration and the training loss approaches zero around the 30th iteration. These results indicate that the BiLSTM-FC network converges faster than the LSTM-FC network in all cases.

With the MRMR algorithm, for the number of features 5 and 15, Table I shows that the classification results with the LSTM-FC network are 87.5% and 95%, respectively, whereas the BiLSTM-FC network achieves classification results of 97.5% in both cases. This result is fully acceptable when compared to the published studies on power system stability classification. For example, the classification accuracy in [11] is 96.3%, in [13] is 96.9%, and in [14] is 97.1%.

A general observation from the study results indicates the superiority of the BiLSTM-FC network over the LSTM-FC network due to its ability to process data in both directions.

VI. CONCLUSION

This paper introduced the implementation of the Minimum Redundancy Maximum Relevance (MRMR) algorithm to reduce the input size for the frequency stability classification model of the Microgrid (MG). The study results on the 16-bus MG system show that the algorithm successfully eliminates redundant variables and selects the important ones. This is of great significance as it helps to reduce the number of measurement sensors, thereby reducing the cost of building the hardware system in the sample collection process. Variable reduction also helps improve classification accuracy and save data storage space.

The paper also presented the application of a deep learning neural network, the Bidirectional Long Short-Term Memory network with Fully Connected layers (BiLSTM-FC), to build a model for classifying the frequency stability of the MG. The study results were obtained on a 16-bus MG system, where the input size was reduced from 89 to only 5, achieving 100% training accuracy and 97.5% testing accuracy. The research results demonstrate that the application of BiLSTM-FC for frequency stability classification in MG is entirely feasible. The findings can be applied to the problem of synthesizing recognition and developing intelligent control strategies to manage grid instability in MG systems.

ACKNOWLEDGMENT

This work belongs to the project T2024-127 funded by Ho Chi Minh City University of Technology and Education, Vietnam.

REFERENCES

[1] J. S. Ali, Y. Qiblawey, A. Alassi, A. M. Massoud, S. M. Muyeen, and H. Abu-Rub, "Power System Stability With High Penetration of Renewable Energy Sources: Challenges, Assessment, and Mitigation Strategies," *IEEE Access*, vol. 13, pp. 39912–39934, 2025, https://doi.org/10.1109/ACCESS.2025.3546491.

[2] M. Debouza, A. Al-Durra, T. H. M. EL-Fouly, and H. H. Zeineldin, "Survey on microgrids with flexible boundaries: Strategies, applications, and future trends," *Electric Power Systems Research*, vol. 205, Apr. 2022, Art. no. 107765, https://doi.org/10.1016/j.epsr.2021.107765.

[3] S. Zarrabian, R. Belkacemi, and A. A. Babalola, "Intelligent mitigation of blackout in real-time microgrids: Neural Network Approach," in *2016 IEEE Power and Energy Conference at Illinois*, Urbana, IL, USA, 2016, pp. 1–6, https://doi.org/10.1109/PECI.2016.7459213.

[4] O. A. Alimi, K. Ouahada, and A. M. Abu-Mahfouz, "A Review of Machine Learning Approaches to Power System Security and Stability," *IEEE Access*, vol. 8, pp. 113512–113531, 2020, https://doi.org/10.1109/ACCESS.2020.3003568.

[5] U. Shahzad, "Artificial Neural Network for Transient Stability Assessment: A Review," in *2024 29th International Conference on Automation and Computing*, Sunderland, United Kingdom, 2024, pp. 1–7, https://doi.org/10.1109/ICAC61394.2024.10718782.

[6] N. Moarref, S. Jafarzadeh, Y. Yaslan, and V. M. Istemihan Genc, "A Transient Stability Prediction Method based on Multi-Channel Convolutional Neural Networks Using Time Series of PMU Measurements," in *2019 11th International Conference on Electrical and Electronics Engineering*, Bursa, Turkey, 2019, pp. 151–155, https://doi.org/10.23919/ELECO47770.2019.8990476.

[7] Z. Li, J. Yan, Y. Liu, W. Liu, L. Li, and H. Qu, "Power system transient voltage vulnerability assessment based on knowledge visualization of CNN," *International Journal of Electrical Power & Energy Systems*, vol. 155, no. B, Jan. 2024, Art. no. 109576, https://doi.org/10.1016/j.ijepes.2023.109576.

[8] X. Li, C. Liu, P. Guo, S. Liu, and J. Ning, "Deep learning-based transient stability assessment framework for large-scale modern power system," *International Journal of Electrical Power & Energy Systems*, vol. 139, Jul. 2022, Art. no. 108010, https://doi.org/10.1016/j.ijepes.2022.108010.

[9] T. Yang, B. Li, and Q. Xun, "LSTM-Attention-Embedding Model-Based Day-Ahead Prediction of Photovoltaic Power Output Using Bayesian Optimization," *IEEE Access*, vol. 7, pp. 171471–171484, 2019, https://doi.org/10.1109/ACCESS.2019.2954290.

[10] Y. Xu, L. Guan, Z. Y. Dong, and K. P. Wong, "Transient stability assessment on China Southern Power Grid system with an improved pattern discovery-based method," in *2010 International Conference on Power System Technology*, Hangzhou, China, 2010, pp. 1–6, https://doi.org/10.1109/POWERCON.2010.5666045.

[11] S. Kalyani and K. S. Swarup, "Pattern analysis and classification for security evaluation in power networks," *International Journal of Electrical Power & Energy Systems*, vol. 44, no. 1, pp. 547–560, Jan. 2013, https://doi.org/10.1016/j.ijepes.2012.07.065.

[12] K. S. Swarup, "Artificial neural network using pattern recognition for security assessment and analysis," *Neurocomputing*, vol. 71, no. 4, pp. 983–998, Jan. 2008, https://doi.org/10.1016/j.neucom.2007.02.017.

[13] R. Zhang, Y. Xu, Z. Y. Dong, and D. J. Hill, "Feature selection for intelligent stability assessment of power systems," in *2012 IEEE Power and Energy Society General Meeting*, San Diego, CA, USA, 2012, pp. 1–7, https://doi.org/10.1109/PESGM.2012.6344780.

[14] N. A. Nguyen, T. N. Le, and H. M. V. Nguyen, "Multi-Goal Feature Selection Function in Binary Particle Swarm Optimization for Power System Stability Classification," *Engineering, Technology & Applied Science Research*, vol. 13, no. 2, pp. 10535–10540, Apr. 2023, https://doi.org/10.48084/etasr.5799.

[15] C. Ding and H. Peng, "Minimum redundancy feature selection from microarray gene expression data," in *Computational Systems Bioinformatics. CSB2003. Proceedings of the 2003 IEEE Bioinformatics Conference. CSB2003*, Stanford, CA, USA, 2003, pp. 523–528, https://doi.org/10.1109/CSB.2003.1227396.

[16] H. Roubhi, A. H. Gharbi, K. Rouabah, and P. Ravier, "Mutual Information-based Feature Selection Strategy for Speech Emotion Recognition using Machine Learning Algorithms Combined with the Voting Rules Method," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 19207–19213, Feb. 2025, https://doi.org/10.48084/etasr.9066.

[17] G. Wang, F. Lauri, and A. H. E. Hassani, "Feature Selection by mRMR Method for Heart Disease Diagnosis," *IEEE Access*, vol. 10, pp. 100786–100796, 2022, https://doi.org/10.1109/ACCESS.2022.3207492.

[18] L. Gong, X. Zhang, T. Chen, and L. Zhang, "Recognition of Disease Genetic Information from Unstructured Text Data Based on BiLSTM-CRF for Molecular Mechanisms," *Security and Communication Networks*, vol. 2021, no. 1, Feb. 2021, Art. no. 6635027, https://doi.org/10.1155/2021/6635027.

[19] Z. Zhao, P. Yang, J. M. Guerrero, Z. Xu, and T. C. Green, "Multiple-Time-Scales Hierarchical Frequency Stability Control Strategy of Medium-Voltage Isolated Microgrid," *IEEE Transactions on Power Electronics*, vol. 31, no. 8, pp. 5974–5991, Aug. 2016, https://doi.org/10.1109/TPEL.2015.2496869.