



A Multidirectional LSTM Model for Predicting the Stability of a Smart Grid

Alazab, Mamoun; Khan, Suleman; Krishnan, Somayaji Siva Rama; Pham, Quoc Viet; Reddy, M. Praveen Kumar; Gadekallu, Thippa Reddy

Published in:
IEEE Access

DOI:
[10.1109/ACCESS.2020.2991067](https://doi.org/10.1109/ACCESS.2020.2991067)

Published: 28/04/2020

Document Version
Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (APA):

Alazab, M., Khan, S., Krishnan, S. S. R., Pham, Q. V., Reddy, M. P. K., & Gadekallu, T. R. (2020). A Multidirectional LSTM Model for Predicting the Stability of a Smart Grid. *IEEE Access*, 8, 85454-85463. Article 9079864. <https://doi.org/10.1109/ACCESS.2020.2991067>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Received April 14, 2020, accepted April 25, 2020, date of publication April 28, 2020, date of current version May 19, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2991067

A Multidirectional LSTM Model for Predicting the Stability of a Smart Grid

MAMOUN ALAZAB¹, (Senior Member, IEEE), SULEMAN KHAN²,
SOMAYAJI SIVA RAMA KRISHNAN³, QUOC-VIET PHAM⁴, (Member, IEEE),
M. PRAVEEN KUMAR REDDY³, AND THIPPA REDDY GADEKALLU³

¹College of Engineering, IT and Environment, Charles Darwin University, Casuarina, NT 0909, Australia

²Air University Islamabad, Islamabad 44000, Pakistan

³School of Information Technology and Engineering, Vellore Institute of Technology, Vellore 632014, India

⁴Research Institute of Computer, Information and Communication, Pusan National University, Busan 46241, South Korea

Corresponding authors: Mamoun Alazab (mamoun.alazab@cdu.edu.au) and Thippa Reddy Gadekallu (thippareddy.g@vit.ac.in)

This work was supported by the Department of Corporate and Information Services, NTG of Australia.

ABSTRACT The grid denotes the electric grid which consists of communication lines, control stations, transformers, and distributors that aids in supplying power from the electrical plant to the consumers. Presently, the electric grid constitutes humongous power production units which generates millions of megawatts of power distributed across several demographic regions. There is a dire need to efficiently manage this power supplied to the various consumer domains such as industries, smart cities, household and organizations. In this regard, a smart grid with intelligent systems is being deployed to cater the dynamic power requirements. A smart grid system follows the Cyber-Physical Systems (CPS) model, in which Information Technology (IT) infrastructure is integrated with physical systems. In the scenario of the smart grid embedded with CPS, the Machine Learning (ML) module is the IT aspect and the power dissipation units are the physical entities. In this research, a novel Multidirectional Long Short-Term Memory (MLSTM) technique is being proposed to predict the stability of the smart grid network. The results obtained are evaluated against other popular Deep Learning approaches such as Gated Recurrent Units (GRU), traditional LSTM and Recurrent Neural Networks (RNN). The experimental results prove that the MLSTM approach outperforms the other ML approaches.

INDEX TERMS Multidirectional long short-term memory (MLSTM), machine learning (ML), smart grid (SG), cyber physical systems (CPS).

I. INTRODUCTION

Cyber-physical system (CPS) is an integration of physical processes, networks, and computation. In CPS, the physical processes are monitored by embedded networks, sensors [1] and computers with the help of a feedback mechanism. The concept of CPS is used in SG, where physical systems like power network infrastructure are integrated with control, intelligence, processing and information. Electricity consumption is increasing day after day with the rise in population. There has been a significant usage of natural resources to generate electricity to cater to this huge demand; however, this method is complex and expensive. There has been extensive research on the grid networks to distribute the power

across various locations effectively. One such approach is the smart grid, which uses Information and Communication Technology (ICT) to aggregate information about the behavior of consumers to create a context-aware system that can distribute the power effectively. Smart grids using Artificial Intelligence (AI) are expected to reduce the need for deploying more power plants for electricity dissipation [2]–[5]. Smart grids also use renewable energy resources to be safely plugged into the grid to appendage the power supply. The intelligent systems, such as Expert Systems (ESs), Fuzzy Logic (FL) [6], [7], Machine Learning (ML) [8], [9] and Deep Neural Networks (DNN) [10]–[12] have revolutionized the power distribution process. These systems render effective tools for design, simulation, fault diagnostics, and fault-tolerant control in the modern smart grid [13]. The recent smart grids for household deployment contain

The associate editor coordinating the review of this manuscript and approving it for publication was Sanjeevikumar Padmanaban¹.

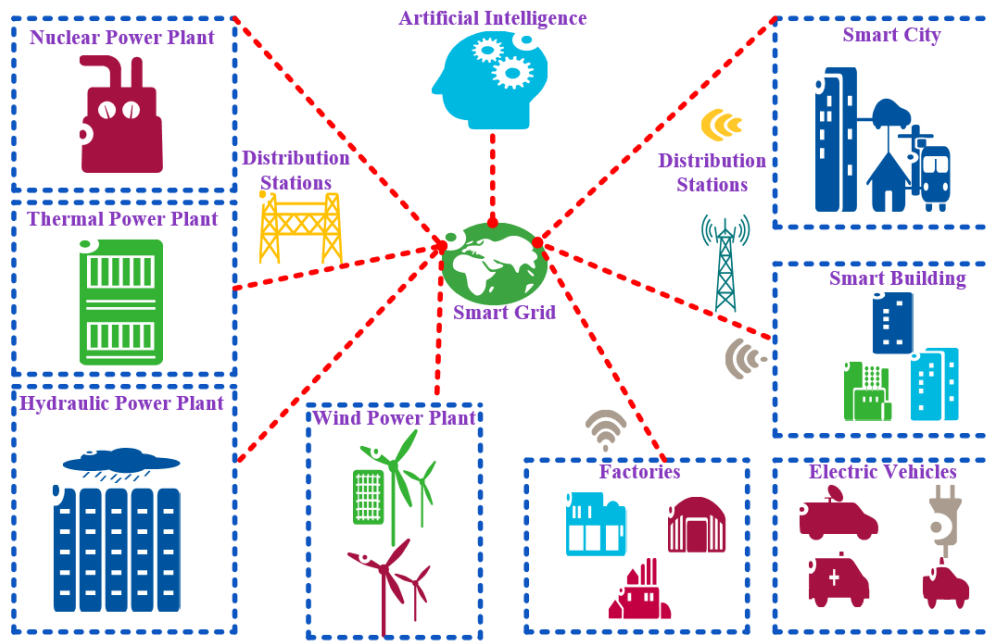


FIGURE 1. Components of the smart grid.

distributed solar-photovoltaic generation and energy storage. This deployment overcomes the decision making issue of using a battery-powered system.

Figure 1 depicts the interaction between the power generation units, distribution centers and other different entities such as industrial factories, electric vehicles, smart buildings, and households. The smart grid plays a significant role in efficiently dissipating the right amount of power to these various entities. The flexibility in the power distribution process is achieved by implementing different AI algorithms in the smart grid.

A smarter grid, which can predict the power demand is the need of the hour. This can be accomplished with the application of Machine Learning (ML) algorithms [14], [15] on the data generated from the grid. The smart grid can help to reduce the pollution and make the electricity price much cheaper.

In this work, a novel Multidirectional Long Short Term Memory (LSTM) model is proposed to predict the stability of SG by classifying the smart grid dataset collected from UCI machine learning repository [16]. The experimental results are then compared with recent deep learning algorithms like Recurrent Neural Networks (RNN), conventional LSTM, and Gated Recurrent Units (GRU). The steps involved in the current work are as follows:

- 1) The Electricity Grid dataset is collected from the UCI machine learning repository.
- 2) The min-max normalization is used for normalizing the dataset.
- 3) Label Encoding technique is used for converting categorical, textual data into numerical data.

- 4) This data is fed to the proposed MLSTM model for training the dataset.
- 5) The proposed model is then evaluated with other deep learning techniques like RNN, GRU, and traditional LSTM using some metrics, including accuracy, precision, recall, and F1-score.

In summary, our contributions in this work can be highlighted as follows:

- A novel MLSTM model is introduced to predict the stability of the smart grid dataset.
- A comprehensive preprocessing is performed on the smart grid dataset.
- An accuracy of 99.07% is achieved which is higher compared to other state-of-the-art deep learning models.

The rest of the paper is organized as follows. Section II discusses recent state-of-the-art literature related to the application of deep learning algorithms on smart grids. In Section III, the proposed model is discussed in detail. Experimental results are discussed in Section IV, which is followed by a conclusion and future work in Section V.

II. LITERATURE SURVEY

In this section recent state-of-the-art works related to application of machine learning techniques on smart grid are discussed.

Technology growth leads to different cybercrimes. One of the most prevalent is, modifying the data in the smart meters [17]. The authors present a framework aggregating Finite Mixture Model clustering for customer segmentation and a Genetic Programming algorithm [18] for recognizing new functions that aid for accurate predictions. The authors

also use a Gradient Boosting Machine algorithm, which is embedded in the framework and performs better than the existing ML algorithms. Most of the existing smart systems use the combination of IoT [19]–[21] embedded with the power of ML algorithms so that the systems perform to their fullest capacity efficiently. The authors in [22] analyze various IoT-based ML modes, which are used in the different domains such as healthcare, smart city, and vehicle-vehicle communication in smart cities. The deployment challenges of various IoT and machine learning systems are also discussed.

A covert data integrity assault (CDIA) on a communications network might be hazardous by directly reducing the availability and safety-measures of a smart grid. This assault is carefully deployed to avoid the conventional bad data detectors in power control stations, and this assault can confront the integrity of the data and instigating a false assessment of the state that would brutally affect the whole power system process. In [23] the authors develop an intelligent system using an unsupervised ML-based model to identify CDIA in smart grids by using non-labeled data. The smart grid is developed using various ICTs, which lead to humongous data originating from various sources. Usage of big data analysis [24] and intelligent systems in smart grid systems helps to solve the challenge of processing and to manage the huge amount of data. Various challenges associated with usage of big data analysis in smart grids are highlighted in [25]. The authors also enlist various applications of big data in smart grid systems.

The estimation of smart grid stability epitomizes a challenging research problem due to the fact that information used for the membership authentication may lead to instabilities in the grid. This is used to govern configurations in which the grid is stable irrespective of abnormalities. The authors in [26] analyze the usage of a machine learning algorithm for forecasting the smart grid stability based on feature extraction. The authors here use three methods for the feature selection process: Binary Particle Swarm optimization Features Selection (BPSOFS), Binary Kangaroo Mob optimization Features Selection (BKMOFS) and Multivariate Adaptive Regression Splines (MARS). The forecast of the grid stability is done by using four classifiers: Logistic Regression (LR), Random Forest (RF), Gradient Boosted Trees (GBT) and Multilayer Perceptron Classifier (MPC).

The age-old electrical grids constituted of one-way communication between electrical grids and consumers. These grids were deployed across the globe, but the efficiency of power management was a significant concern. To solve this challenge, the smart grids evolved with two-way communication between the grid and the consumer. The primary objective behind developing this smart grid was to accurately predict the energy requirements in a given demographic. The parameters such as historical weather, load, and energy generation data [27] are used to implement machine learning algorithms for predictions. The researchers also develop two models, namely: linear regression model, and deep neural

network [28]. These models were assessed by means of root mean squared error and it was observed that the deep neural network models out-performed the linear regression models in forecasting load and energy production at a given region.

Energy load prediction in the smart grid is forecasting the electrical power requirements to meet dynamic demands. There is a dire need to embed accuracy for load prediction which aids electrical services in monitoring their energy generation and process. Presently most of the prediction in any system is done through machine learning algorithms to achieve maximum efficacy. Apache Spark and Apache Hadoop [29] are proposed as big data frameworks for distributed processing for prediction of the energy load. The authors also use MLlib to assess the prediction accuracy of various regression methods such as linear regression, generalized linear regression, decision tree, random forest and gradient-boosted trees.

Currently, the smart grids are in trend for the deployment of power systems to manage the energy dissipation efficiently. Since the deployment of a smart grid network involves huge complexities due to the humongous data being generated, the use of artificial intelligent systems helps in easing this process. The advancements in intelligent systems evolution has paved a way for better methods such as Deep Learning (DL), Reinforcement Learning (RL) and deep reinforcement learning (DRL). The challenges in deploying these methods in Smart grid are enlisted to enable further investigation by future researchers [30].

The dynamic energy consumption in household appliances is a major concern for sustainability and efficacy in building smart cities. The advancement in IoT technologies enables newer energy management techniques to address this dynamic energy usage. The authors in [31] suggest an implementation of probabilistic data-driven prognostic technique for consumption prediction in inhabited constructions. This technique uses a Bayesian Network (BN) framework which enables the system to determine dependency relations amongst contributing variables. The authors assess the proposed technique by using the datasets given by Pacific Northwest National Lab (PNNL) which was aggregated through a pilot Smart Grid project.

Various big giants have come together, forming a consortium to address the challenges and future needs of smart grids. Smart grids use different Artificial Intelligence (AI) methods such as Artificial Neural Network (ANN), Machine Learning (ML) and Deep Learning (DL) for efficient energy consumption. Different algorithms of DL for load prediction issues are proposed in [32] for the smart grid. The authors focus on the usage of different apps of DL for load prediction in the smart grid network. The authors also compare the accuracy results of Root Mean Square Error and Mean Absolute Error for the studied applications and their results infer that the use of convolutional neural network with k-means algorithm had a huge percentage of reduction in case of root mean square error.

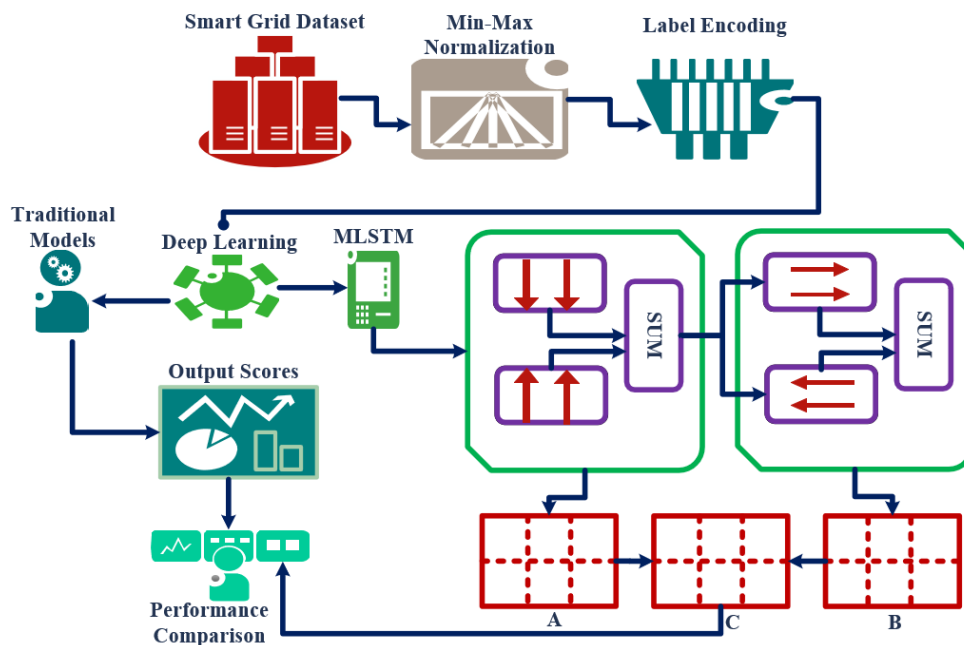


FIGURE 2. Workflow of the proposed multi-directional LSTM model.

Electricity price has a significant impact on the stability of the distributed power grids. The cost sensitivity and reaction times of power producers and consumers also influence the stability factor. Wood [33] propose a model called decentral smart grid control (DSGC) to render the demand-side control of distributed power grids by associating the electricity price to variations in grid frequency upon the time gauge of a few seconds. The authors simulate the power demand-side consumption/production on analogous time gauges. The authors also implement an optimized data-matching machine-learning algorithm, the transparent, open box (TOB) learning network to envisage dynamic grid stability for the simulation from its independent variables. Training the vast data with variations in the dimensionality reduces the effectiveness of the machine learning model. The authors in [34] implement a secondary principal component analysis (PCA) algorithm [35] to decrease the data dimensions. This algorithm is applied to manage the ML techniques to increase the stability of the grid systems.

From the above discussion, there is no foolproof system to address the challenges of stability in a smart grid network. In this work, a novel Multi-Directional LSTM (MLSTM) architecture to train the SG dataset is proposed.

III. PROPOSED MULTI-DIMENSIONAL LSTM MODEL

The workflow of the proposed MSLTM model to analyze the power utilization [36], [37] and predict the stability of the SG dataset is depicted in Figure 2. First, the electricity grid dataset from different power generating units is aggregated. The dataset is then normalized using min-max normalization. During this process, the minimum and maximum values of

the data are obtained and replaced by using Eq. (1).

$$\bar{L} = \frac{l - \min(X)}{\max(X) - \min(X)}(new_min(X) - new_max(X)) + new_min(X), \quad (1)$$

where X represents an attribute in the given dataset, $\min(X)$ and $\max(X)$ denotes minimum and maximum values of the attributes in the given dataset, respectively. \bar{L} indicates the updated value of every entry in a dataset, l corresponds to the previous values in the dataset, $new_max(X)$ and $new_min(X)$ denotes the upper and lower boundary values in the given range, respectively.

The ML algorithms cannot process the categorical values in the dataset. This is the reason behind using the Label encoding technique which converts the categorical values in the dataset into numerical values suitable for processing by the ML algorithms. In the next step the smart grid dataset is trained with the proposed MLSTM approach. The performance of the proposed model is then compared with RNN, traditional LSTM, and GRU models by using metrics such as accuracy, precision, recall, and F1-score. These models are explained in brief in the following.

Recurrent Neural Networks (RNN): Traditional Neural Networks (NN) do not yield satisfactory results on time series data due to the vanishing gradient problem. In order to handle the issue mentioned above, RNN was introduced in 1982 by Hopfield [38]. RNN has the advantage of making the NNs learn the patterns over a period of time. RNN can predict the sequential data like actions based on previous events in a video, audio from speech, events from the text, etc. The working model of RNN is shown in Figure 3. In the Figure, X_t

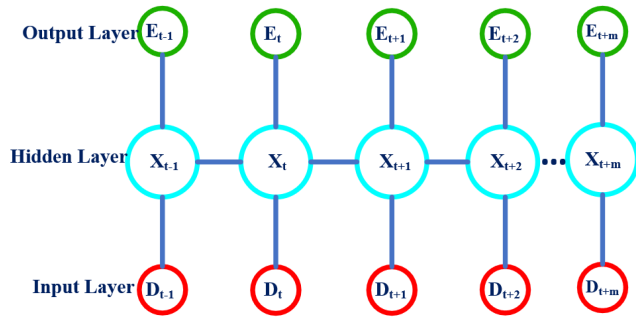


FIGURE 3. Working model of RNN.

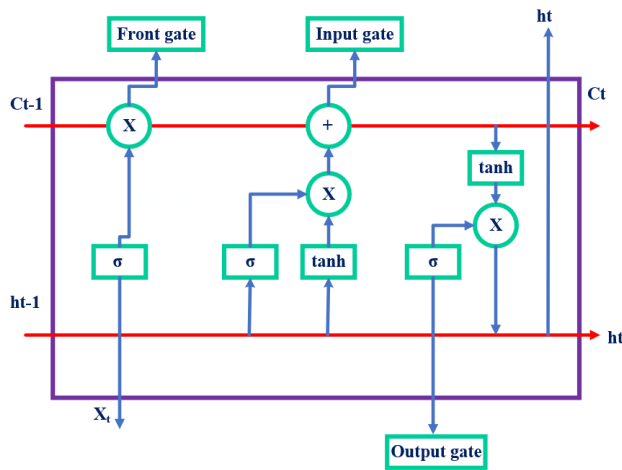


FIGURE 4. Long Short Term Memory Architecture.

represents the hidden layer's weight vector, E_t represents the output layer's weight vector, D_t denotes input word vector. The hidden layer time stamp t is calculated by using Eq. (2).

$$X_t = \sigma(A * D_t + C * X_{t-1}), \quad (2)$$

$$E_t = \sigma(B * X_t), \quad (3)$$

where $\sigma(\cdot)$ is considered as the activation function. The activation function can be either Sigmoid, Tanh, or Relu. At every timestamp t , the hidden state X_t is computed by using Eq. (2) with the corresponding inputs and parameters.

Long Short Term Memory (LSTM): RNN uses more number of network layers, which is a challenge to gain an insight about the parameters from previous layers. In recent studies, the LSTM network is one of the popular approaches to overcome this challenge [39]. LSTM has a chain structure like RNN that has multiple neural network modules. Figure 4 illustrates the architecture of LSTM, which consists of different gates like input gate, output gate and forget gate. These gates select and reject the information passing through the network.

Input gate $i^{(t)}$ consists of \tanh as activation function ranging from -1 to 1 . It takes the current input $x^{(t)}$, parameters $C^{(t-1)}$ and $h^{(t-1)}$ for processing. Forget gate $f^{(t)}$ has sigmoid and \tanh as the activation function. The forget gate decides

how much of the information from the previous output has to be retained. If the value is 1 , then the data will be transformed into the network, if it is 0 , the data will not be passed through the network. The output gate $o^{(t)}$ has sigmoid as an activation function ranging from -1 to 1 . At every timestamp, $i^{(t)}$, $o^{(t)}$, $f^{(t)}$ are computed by using the following equations.

$$i^{(t)} = \sigma(W^i[C^{t-1}, h^{(t-1)}, x^{(t)}] + b^i), \quad (4)$$

$$o^{(t)} = \sigma(W^o[C^{t-1}, h^{(t-1)}, x^{(t)}] + b^o), \quad (5)$$

$$f^{(t)} = \sigma(W^f[C^{t-1}, h^{(t-1)}, x^{(t)}] + b^f). \quad (6)$$

Traditional LSTM works bi-directionally, while in this research, two LSTMs are used; one LSTM for upward and downward scanning and the other LSTM is for right and left scanning. The input of the second LSTM is a summation of the first LSTM. The proposed MLSTM utilizes twice the number of input gates, output gates and forget gates as compared to the traditional LSTM. MLSTM achieves a promising accuracy, but it has a computational overhead.

Gated Recurrent Unit (GRU): GRU is the latest variant of RNN designed to deal with short-term memory problems similar to LSTM. GRU does not have a cell state and makes use of a hidden state to carry information. GRU also consists of two gates: a reset gate r^t and an update gate z^t , which are represented in the following equations:

$$z^t = \sigma(w_{z^t}x^t + U_{z^t}h_{t-1} + b_{z^t}), \quad (7)$$

$$r^t = \sigma(w_{r^t}x^t + U_{r^t}h_{t-1} + b_{r^t}), \quad (8)$$

where z^t denotes update gate, $\sigma(\cdot)$ represents the sigmoid function, w , U and b are parameter matrices and vector, h_t denotes the output vector, x^t denotes the input vector. The update gate performs similar functions of the forget gate and an input gate of an LSTM. It is responsible for deciding which information has to be dropped and which information has to be added. The reset gate is responsible for determining the amount of the previous data to be forgotten. Since GRU has fewer gates compared to LSTM, the training process is normally faster.

A. MLSTM MODEL

The proposed model consists of four 1D spatial LSTMs to scan each column and row in different directions independently. At every step, hidden layers are calculated and summed at the end. Two spatial 1D LSTM are applied in (A) vertically, where the hidden states are calculated at every step and the results are summed. Similarly, in (B), two spatial 1D LSTM are applied in horizontal directions. The combination of A and B defines the Multidirectional LSTM, as shown in Figure 5.

As presented in Figure 5, the suggested module is a combination of the 1D spatial LSTM (A) and (B). In each module, two 1D LSTMs are used for scanning across the attributes maps vertically or horizontally in both the directions (bi-directional LSTMs) and their hidden states are updated at every spatial step. For each of the vertical and horizontal directions,

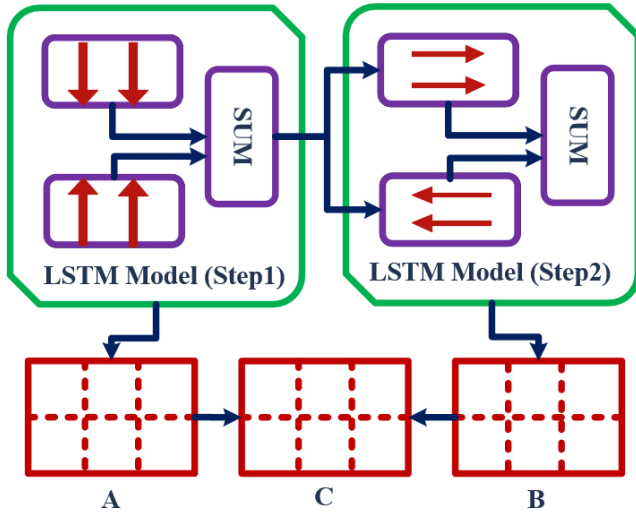


FIGURE 5. The general architecture of the proposed multi-directional LSTM.

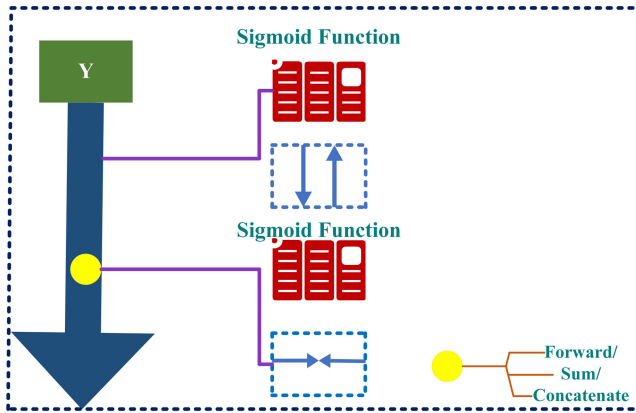


FIGURE 6. 1D MLSTM Sigmoid Function Working.

two output attribute maps are used. In the current simulation, a sum of these output attribute maps is used (an alternative is to concatenate the output feature maps, but that would increase the number of parameters).

For illustration, we consider the input data for the 1D spatial LSTM. F^L is an attribute vector at every spatial location, every column or row on the function cards is treated as a series. While scanning from top to bottom, the attribute response for m, n represent the dimensionality of features and it can be estimated as:

$$F_{m,n}^{L+1} = A \left(U F_{m,n}^L + V F_{m,n-1}^{L+1} + b \right), \quad (9)$$

where $F_{m,0}^{L+1} = 0$, $F_{m,n}^L \in R^{d \times 1}$, $F_{m,n}^{L+1}$, $F_{m,n-1}^{L+1} \in R^{D \times 1}$, $U \in R^{D \times d}$, $V \in R^{D \times D}$, $b \in R^{D \times 1}$, D depicts the number of nodes used in 1D Spatial LSTM and F represents the non-linearity function. Similarly, right to left 1D spatial LSTM can also be calculated using the same approach. Here d represents number of attributes in the dataset, while $*$ depicts the multiplication.

Figure 6 represents the working of the proposed model at every hidden layer. To process the data at hidden layers, the MLSTM model uses three approaches, namely forward propagation, sum, and concatenation as follows:

$$Y^{(L+1)} = F(Y^L, W), \quad (10)$$

$$Y^{(L+1)} = Y^L + F(Y^L, W), \quad (11)$$

$$Y^{(L+1)} = [Y^L; F(Y^L, W)]. \quad (12)$$

IV. RESULTS AND DISCUSSION

This section discusses about prediction of smart grid stability using the proposed model.

A. EXPERIMENTAL SETUP

The experimentation is carried out using an online Graphical Processing Unit (GPU) provided by Google called “Google Colab”. A personal computer with Windows 8.1 Operating system and core I3 processor is used. The programming language used for this purpose is Python 3.7.

B. DATASET DESCRIPTION

The dataset used for the experimentation is collected from the UCI machine learning repository [16]. The dataset has 10000 instances with 14 attributes. The attributes in the dataset have information related to the electricity producer values, nominal power consumed/produced, coefficient related to price elasticity, the maximum value of the equation root, the stability of the system (class label, whether the system is stable or not).

C. METRICS USED FOR EVALUATING THE PROPOSED APPROACH

The performance metrics used to evaluate the proposed MLSTM model are explained below.

Confusion matrix: A confusion matrix is used to evaluate the performance of a classification model.

- **Accuracy:** Accuracy signifies the correctness of a classifier. Accuracy is calculated as follows:

$$Accuracy = \frac{(True\ Positive + True\ Negative)}{Total\ Instances}. \quad (13)$$

- **Precision:**

$$Precision = \frac{True\ Positive}{(Predicted\ Instances = True)}. \quad (14)$$

- **Recall:**

$$Recall = \frac{True\ Positive}{Actual\ number\ of\ instances\ as\ True}. \quad (15)$$

- **F1 Score:**

$$F1\ Measure = \frac{2 * Precision * Recall}{Precision + Recall} \quad (16)$$

TABLE 1. Confusion Matrix for DNN Models.

RNN Confusion Matrix			GRU Confusion Matrix			LSTM Confusion Matrix			Propose Model		
Labels	Stable	Unstable	Labels	Stable	Unstable	Labels	Stable	Unstable	Labels	Stable	Unstable
Stable	970	114	Stable	1081	3	Stable	1013	11	Stable	1084	0
Unstable	5	1911	Unstable	99	1817	Unstable	75	1841	Unstable	28	1888

TABLE 2. Average Training and Testing Accuracy and Loss For Deep Learning Models.

verage Training and Testing Accuracy and Loss For Deep Learning Models.

Classifier	Training Accuracy %	Training Loss	Testing Accuracy %	Testing Loss
GRU	97.17	0.06	97.30	0.06
RNN	96.66	0.08	96.60	0.08
LSTM	97.30	0.06	97.13	0.06
Proposed Model	99.07	0.02	99.07	0.02

D. PERFORMANCE EVALUATION OF THE PROPOSED MODEL

70% of the SG dataset is used for training and validation, and the remaining 30% of the dataset is used for testing and validation purpose.

Various evaluation methods are used for experimentation, which include true positive, true negative, false negative, false positive, precision, recall, f1 score, training loss and testing loss. At the end, Receiver Operating Characteristic (ROC) curve is also used to justify the results. The proposed MLSTM model is found to be very effective for the smart grid stability system. The accuracy for both stable and unstable labels are more than 99%. Table 1 depicts the confusion matrix for traditional deep learning models and proposed MLSTM model. Proposed MLSTM detects 1084 data as true positives and 1888 true negatives accurately. Only 28 data are misclassified by the proposed model. Training and testing accuracy achieved using GRU is 97.17% and 97.30% respectively. Similarly, for training and testing the data loss for GRU is 0.06. Using RNN the training and testing, accuracy is recorded as 96.66% and 96.60% respectively. The data loss incurred is 0.08 and 0.08 for both training and testing using RNN. Proposed MLSTM achieved 99.07% accuracy and 0.02 loss for both training and testing respectively. The proposed method achieved 2% higher accuracy when compared to GRU, RNN and LSTM as shown in Table 2. The proposed MLSTM achieved 97% and 100% precision for stable and unstable class respectively. Recall for both stable and unstable classes is recorded as 100% and 99% respectively. F1-Scores for detecting the stability of the smart grid systems for stable and unstable classes is 99% respectively. ROC for MLSTM is 99.27% which is higher in comparison with GRU and LSTM.

From Figure 7 it can be concluded that the proposed model achieved 0.02 loss while other traditional deep learning models GRU and LSTM achieved 0.06 loss respectively. Similarly RNN achieved 0.08 loss which is high when compared to proposed model. From Figure 7, it is evident that the proposed model performs well as the loss incurred in minimum when compared to other traditional deep learning models.

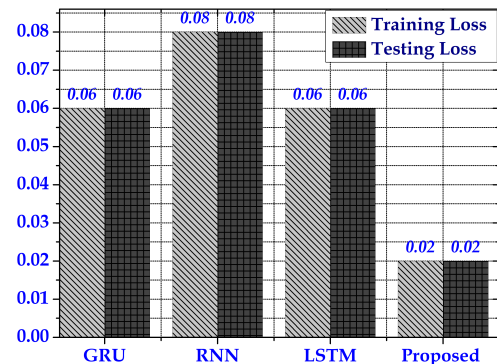
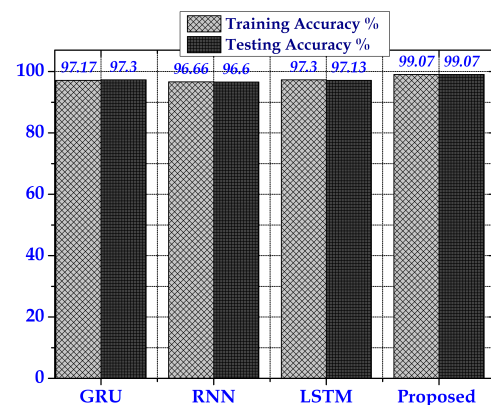
**FIGURE 7. Training and Testing Loss of All The Models.****FIGURE 8. Training and Testing Accuracies of All The Models.**

Figure 8 depicts that the proposed model achieved 99.07% accuracy for training and testing respectively as compared to the other models. GRU achieved 97.17% and 97.30% accuracy for training and testing respectively. While RNN achieved 96.66% accuracy for training and 96.60% accuracy for testing respectively. Similarly LSTM achieved 97.30% accuracy for training and 97.13% accuracy for testing respectively. The proposed model achieved 3% increase in training and testing accuracy.

Figure 9, 10, 11 and Figure 12 depicts that the proposed model achieved high area under the curve (AUC) which is 99.27%. When compared to RNN, GRU and LSTM respectively the proposed model achieved 2% higher AUC.

Table 3 and Figure 13 shows the effectiveness of the proposed model in terms of precision, recall and f1-score. The proposed model achieved 97% precision while other deep learning models like GRU, RNN and LSTM achieved

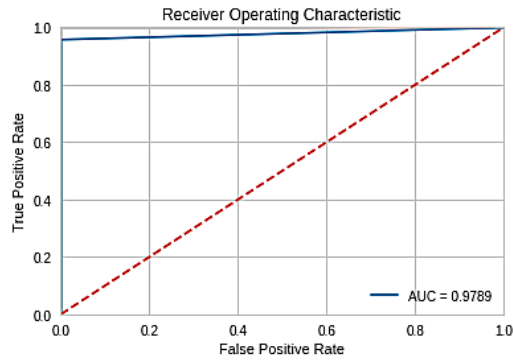


FIGURE 9. ROC Curve for GRU.

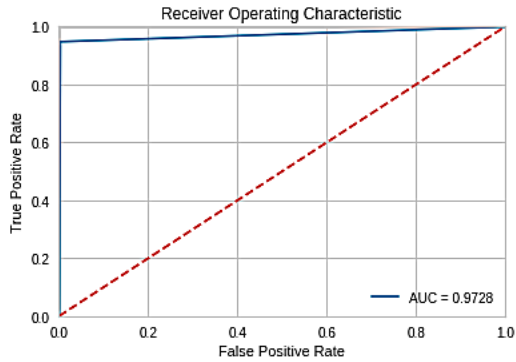


FIGURE 10. ROC Curve for RNN.

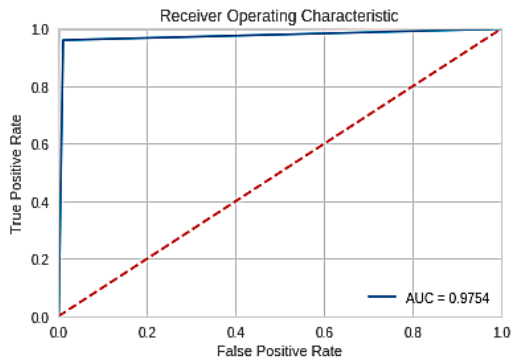


FIGURE 11. ROC Curve for LSTM.

TABLE 3. Classification Report For Deep Learning Models.

Classifier	Label	Precision %	Recall %	F1 Score %
GRU	Stable	93.00	100.00	96.00
	Unstable	100.00	96.00	98.00
RNN	Stable	92.00	100.00	95.00
	Unstable	100.00	95.00	97.00
LSTM	Stable	93.00	99.00	96.00
	Unstable	99.00	96.00	98.00
Proposed Model	Stable	97.00	100.00	99.00
	Unstable	100.00	99.00	99.00

93.00%, 92.00% and 93.00% precision for stable class respectively. The proposed model achieved 4% more precision for stable class compare to other deep learning models.

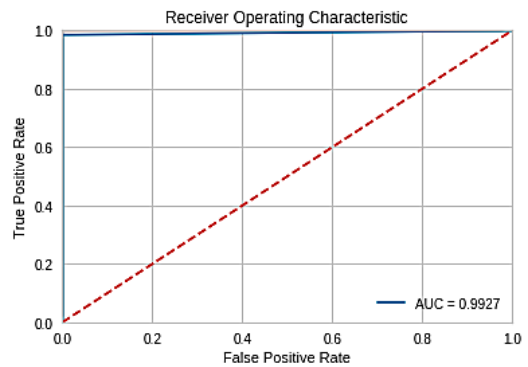


FIGURE 12. ROC Curve for Proposed Model.

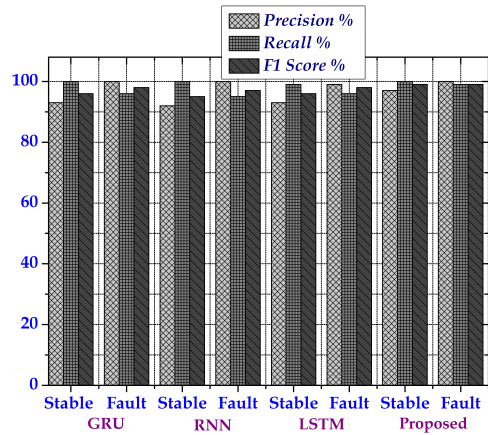


FIGURE 13. Classification Report of All The Models.

Similarly F1-Score for GRU is 96.00%, RNN achieved 97.00% and for LSTM f1-score was 96.00%. The proposed model achieved 99.00% F1-score which is also higher when compared to mentioned deep learning models for stable class. Similarly, for unstable class the proposed model outperformed the traditional models in terms of precision, recall and F1-score.

V. CONCLUSION

Smart grid is an application of cyber physical system which is used to intelligently manage power dissipation across different entities. The stability of the smart grid is essential for efficient power distribution to the control stations. Machine learning techniques play a vital role in predicting the stability of the smart grid. In this work a novel MLSTM model is introduced to predict the stability of smart grid. The proposed model is experimented on the smart grid dataset from UCI Machine Learning Repository. The performance of MLSTM is compared with traditional ML models like LSTM, GRU, RNN. The comparative analysis proves the superiority of the proposed model with respect to accuracy, precision, loss and ROC curve metrics. The proposed model achieved 99.07% training and testing accuracy which is 3% times higher compared to other traditional deep learning

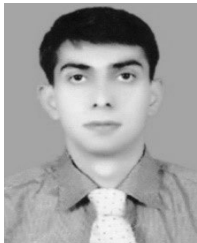
models. Training and testing loss for the proposed model is 0.02 where as for GRU, RNN and LSTM it is 0.06, 0.08 and 0.06 respectively. Precision, recall and F1-score for stable class is 97.00%, 100.00% and 99.00% respectively for the proposed model which is 3% higher compared to other traditional deep learning models. Similarly for unstable class the proposed model achieved 100.00% precision, 99.00% recall and 99.00% F1-score respectively. ROC curve also highlights the effectiveness of the proposed model. 99.27% ROC is achieved by the proposed model which outperformed the other models. As part of the future work context aware model can be deployed to cater the dynamic power requirements and also to make the smart grids more reliable.

REFERENCES

- [1] M. Numan, F. Subhan, W. Z. Khan, S. Hakak, S. Haider, G. T. Reddy, A. Jolfaei, and M. Alazab, "A systematic review on clone node detection in static wireless sensor networks," *IEEE Access*, vol. 8, pp. 65450–65461, 2020.
- [2] M. L. Tuballa and M. L. Abundo, "A review of the development of smart grid technologies," *Renew. Sustain. Energy Rev.*, vol. 59, pp. 710–725, Jun. 2016.
- [3] I. F. Siddiqui, S. U.-J. Lee, A. Abbas, and A. K. Bashir, "Optimizing lifespan and energy consumption by smart meters in green-cloud-based smart grids," *IEEE Access*, vol. 5, pp. 20934–20945, 2017.
- [4] K. Mahmud, A. K. Sahoo, E. Fernandez, P. Sanjeevikumar, and J. B. Holm-Nielsen, "Computational tools for modeling and analysis of power generation and transmission systems of the smart grid," *IEEE Syst. J.*, early access, Jan. 27, 2020, doi: 10.1109/JSYST.2020.2964436.
- [5] N. Kumar, S. Zeadally, and J. J. P. C. Rodrigues, "Vehicular delay-tolerant networks for smart grid data management using mobile edge computing," *IEEE Commun. Mag.*, vol. 54, no. 10, pp. 60–66, Oct. 2016.
- [6] G. T. Reddy, M. P. K. Reddy, K. Lakshmana, D. S. Rajput, R. Kaluri, and G. Srivastava, "Hybrid genetic algorithm and a fuzzy logic classifier for heart disease diagnosis," *Evol. Intell.*, pp. 1–12, Nov. 2019.
- [7] N. Khare and G. T. Reddy, "Heart disease classification system using optimised fuzzy rule based algorithm," *Int. J. Biomed. Eng. Technol.*, vol. 27, no. 3, pp. 183–202, 2018.
- [8] M. Shafiq, X. Yu, A. K. Bashir, H. N. Chaudhry, and D. Wang, "A machine learning approach for feature selection traffic classification using security analysis," *J. Supercomput.*, vol. 74, no. 10, pp. 4867–4892, Oct. 2018.
- [9] C. Iwendi, S. Khan, J. H. Anajemba, A. K. Bashir, and F. Noor, "Realizing an efficient IoMT-assisted patient diet recommendation system through machine learning model," *IEEE Access*, vol. 8, pp. 28462–28474, 2020.
- [10] O. Irshad, M. U. G. Khan, R. Iqbal, S. Basheer, and A. K. Bashir, "Performance optimization of IoT based biological systems using deep learning," *Comput. Commun.*, vol. 155, pp. 24–31, Apr. 2020.
- [11] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, and S. Venkatraman, "Robust intelligent malware detection using deep learning," *IEEE Access*, vol. 7, pp. 46717–46738, 2019.
- [12] M. Alazab and M. Tang, *Deep Learning Applications for Cyber Security*. Springer, 2019.
- [13] B. K. Bose, "Artificial intelligence techniques in smart grid and renewable energy systems—Some example applications," *Proc. IEEE*, vol. 105, no. 11, pp. 2262–2273, Oct. 2017.
- [14] T. R. Gadekallu, N. Khare, S. Bhattacharya, S. Singh, P. K. R. Maddikunta, I.-H. Ra, and M. Alazab, "Early detection of diabetic retinopathy using PCA-firefly based deep learning model," *Electronics*, vol. 9, no. 2, p. 274, 2020.
- [15] S. Bhattacharya, S. R. K. S., P. K. R. Maddikunta, R. Kaluri, S. Singh, T. R. Gadekallu, M. Alazab, and U. Tariq, "A novel PCA-firefly based XGBoost classification model for intrusion detection in networks using GPU," *Electronics*, vol. 9, no. 2, p. 219, 2020.
- [16] V. Arzamasov, K. Bohm, and P. Jochem, "Towards concise models of grid stability," in *Proc. IEEE Int. Conf. Commun., Control, Comput. Technol. Smart Grids (SmartGridComm)*, Oct. 2018, pp. 1–6.
- [17] A. Jindal, A. Dua, K. Kaur, M. Singh, N. Kumar, and S. Mishra, "Decision tree and SVM-based data analytics for theft detection in smart grid," *IEEE Trans. Ind. Informat.*, vol. 12, no. 3, pp. 1005–1016, Jun. 2016.
- [18] R. Razavi, A. Gharipour, M. Fleury, and I. J. Akpan, "A practical feature-engineering framework for electricity theft detection in smart grids," *Appl. Energy*, vol. 238, pp. 481–494, Mar. 2019.
- [19] Y. Koth, I. Al Ridhawi, M. Aloqaily, T. Baker, Y. Jararweh, and H. Tawfik, "Cloud-based multi-agent cooperation for IoT devices using workflow-nets," *J. Grid Comput.*, vol. 17, no. 4, pp. 625–650, Dec. 2019.
- [20] Q.-V. Pham, F. Fang, V. N. Ha, M. Le, Z. Ding, L. B. Le, and W.-J. Hwang, "A survey of multi-access edge computing in 5G and beyond: Fundamentals, technology integration, and state-of-the-art," *CoRR*, vol. abs/1906.08452, 2019. [Online]. Available: <http://arxiv.org/abs/1906.08452>
- [21] C. Iwendi, P. K. R. Maddikunta, T. R. Gadekallu, K. Lakshmana, A. K. Bashir, and M. J. Piran, "A Metaheuristic optimization approach for energy efficiency in the IoT networks," *Softw., Pract. Exper.*, Feb. 2020.
- [22] I. U. Din, M. Guizani, J. J. P. C. Rodrigues, S. Hassan, and V. V. Korotaev, "Machine learning in the Internet of Things: Designed techniques for smart cities," *Future Gener. Comput. Syst.*, vol. 100, pp. 826–843, Nov. 2019.
- [23] S. Ahmed, Y. Lee, S.-H. Hyun, and I. Koo, "Unsupervised machine learning-based detection of covert data integrity assault in smart grid networks utilizing isolation forest," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 10, pp. 2765–2777, Oct. 2019.
- [24] S. Zerdoumi, A. Q. M. Sabri, A. Kamsin, I. A. T. Hashem, A. Gani, S. Hakak, M. A. Al-garadi, and V. Chang, "Image pattern recognition in big data: Taxonomy and open challenges: Survey," *Multimedia Tools Appl.*, vol. 77, no. 8, pp. 10091–10121, Apr. 2018.
- [25] M. Ghorbanian, S. H. Dolatabadi, and P. Siano, "Big data issues in smart grids: A survey," *IEEE Syst. J.*, vol. 13, no. 4, pp. 4158–4168, Dec. 2019.
- [26] D. Moldovan and I. Salomie, "Detection of sources of instability in smart grids using machine learning techniques," in *Proc. IEEE 15th Int. Conf. Intell. Comput. Commun. Process. (ICCP)*, Sep. 2019, pp. 175–182.
- [27] C. Chu, "Using machine learning to predict renewable energy generation and smart grid designs," M.S. thesis, Princeton Univ., 2019.
- [28] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep learning approach for intelligent intrusion detection system," *IEEE Access*, vol. 7, pp. 41525–41550, 2019.
- [29] D. Syed, S. S. Refaat, and H. Abu-Rub, "Performance evaluation of distributed machine learning for load forecasting in smart grids," in *Proc. Cybern. Informat. (KI)*, Jan. 2020, pp. 1–6.
- [30] D. Zhang, X. Han, and C. Deng, "Review on the research and practice of deep learning and reinforcement learning in smart grids," *CSEE J. Power Energy Syst.*, vol. 4, no. 3, pp. 362–370, Sep. 2018.
- [31] N. Bassamzadeh and R. Ghanem, "Multiscale stochastic prediction of electricity demand in smart grids using Bayesian networks," *Appl. Energy*, vol. 193, pp. 369–380, May 2017.
- [32] A. Almalaq and G. Edwards, "A review of deep learning methods applied on load forecasting," in *Proc. 16th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2017, pp. 511–516.
- [33] D. A. Wood, "Predicting stability of a decentralized power grid linking electricity price formulation to grid frequency applying an optimized data-matching learning network to simulated data," *Technol. Econ. Smart Grids Sustain. Energy*, vol. 5, no. 1, pp. 1–21, Dec. 2020.
- [34] K. Chen, "Indirect PCA dimensionality reduction based machine learning algorithms for power system transient stability assessment," in *Proc. IEEE Innov. Smart Grid Technol. Asia (ISGT Asia)*, May 2019, pp. 4175–4179.
- [35] G. T. Reddy, M. P. K. Reddy, K. Lakshmana, R. Kaluri, D. S. Rajput, G. Srivastava, and T. Baker, "Analysis of dimensionality reduction techniques on big data," *IEEE Access*, vol. 8, pp. 54776–54788, 2020.
- [36] E. Hossain, M. R. Tur, S. Padmanaban, S. Ay, and I. Khan, "Analysis and mitigation of power quality issues in distributed generation systems using custom power devices," *IEEE Access*, vol. 6, pp. 16816–16833, 2018.
- [37] K. Krishnamurthy, S. Padmanaban, F. Blaabjerg, R. B. Neelakandan, and K. R. Prabhu, "Power electronic converter configurations integration with hybrid energy sources—A comprehensive review for state-of the-art in research," *Electr. Power Compon. Syst.*, vol. 47, no. 18, pp. 1623–1650, Nov. 2019.
- [38] J. Hopfield, "Hopfield network," *Scholarpedia*, vol. 2, no. 5, p. 1977, 2007.
- [39] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger, Eds. Curran Associates, 2014, pp. 3104–3112.



MAMOUN ALAZAB (Senior Member, IEEE) received the Ph.D. degree in computer science from the School of Science, Information Technology and Engineering, Federation University of Australia. He is currently an Associate Professor with the College of Engineering, IT, and Environment, Charles Darwin University, Australia. He is also a Cyber Security Researcher and a Practitioner with industry and academic experience. His research is multidisciplinary that focuses on cyber security and digital forensics of computer systems with a focus on cyber-crime detection and prevention. He has more than 150 research articles in many international journals and conferences, he delivered many invited and keynote speeches, 24 events in 2019 alone. He convened and chaired more than 50 conferences and workshops. He also works closely with government and industry on many projects, including Northern Territory (NT) Department of Information and Corporate Services, IBM, Trend Micro, the Australian Federal Police (AFP), the Australian Communications and Media Authority (ACMA), Westpac, and United Nations Office on Drugs and Crime (UNODC). He is also the Founding Chair of the IEEE NT Subsection.



SULEMAN KHAN received the master's degree from the Department of Computer Science, Air University, Islamabad, Pakistan, in 2019. He is currently a Research Associate with Air University. His research interests include network security, machine learning, and data science.



SOMAYAJI SIVA RAMA KRISHNAN is currently a Research Member with the Centre for Ambient Intelligence and Advanced Networking Research (AMIR) and also working with VIT, as an Assistant Professor Senior. He has working experience with the Centre for Development and Advance Computing (C-DAC) (Ministry of Science and Technology, Government of India), as a Research Intern in the area of data center technologies. He is also certified by EMC Corp., as a Proven Professional in information storage and management. He is also an EMC Academic Alliance Faculty and played a key role in establishing a MoU between VIT University and EMC Corp. He had proposed and developed an Intelligent Network design framework for building small and large scale network. He also developed an efficient and secure framework for IP Storage network for C-DAC. His current interests include e-waste management in India, wireless networks, and cloud computing.



QUOC-VIET PHAM (Member, IEEE) received the B.S. degree in electronics and telecommunications engineering from the Hanoi University of Science and Technology, Vietnam, in 2013, and the M.S. and Ph.D. degrees in telecommunications engineering from Inje University, South Korea, in 2015 and 2017, respectively. He was a Research Fellow with the ICT Convergence Center, Changwon National University, South Korea. He is currently a Research Professor with the Research Institute of Computer, Information, and Communication, Pusan National University, South Korea. His research interests include convex optimization, game theory, and machine learning to analyze, design, and optimize resource allocation in 5G wireless networks and beyond. He received the Best Ph.D. Dissertation Award in engineering from Inje University, in 2017.



M. PRAVEEN KUMAR REDDY received the B.Tech. degree in CSE from JNT University, and the M.Tech. degree in C.S.E. and the Ph.D. degree from VIT, Vellore, India. He was a Visiting Professor with the Guangdong University of Technology, China, in 2019. He had worked as a Software Developer with IBM, in 2011. He worked with Alcatel-Lucent, in 2013. He produced more than 15 international/national publications. He is currently working as an Assistant Professor with the School of Information Technology and Engineering, VIT. He is also working in the area of energy aware applications for Internet of Things (IoT) and high-performance computing.



THIPPA REDDY GADEKALLU received the B.Tech. degree in CSE from Nagarjuna University, the M.Tech. degree in CSE from Anna University, Chennai, India, and the Ph.D. degree from VIT, Vellore, India. He is currently working as an Assistant Professor (Senior) with the School of Information Technology and Engineering, VIT. He has 14 years of experience in teaching. He produced more than 25 international/national publications. He is also working in the area of machine learning, deep neural networks, the Internet of Things, and blockchain.

...