An Investigation into the Methods and Applications of Deep Learning in Smart Grid

Rituraj Rituraj
An Investigation into the Methods and Applications of Deep Learning in Smart Grid

Rituraj Rituraj
Doctoral School of Applied Informatics and Applied Mathematics
Obuda University
Budapest, Hungary
rituraj88@stud.uni-obuda.hu

Abstract—The manuscript represents the state-of-the-art review of the deep learning methods for smart grid applications. This paper reviews novel applications of deep learning algorithms in smart grid. The deep learning based three algorithms i.e., Long-Short Term Memory, recurrent neural network, convolution neural network found to be most useful in smart grid application. These algorithms are found to be most useful for forecasting, cyber-attack, anomaly detection and electricity theft in smart grid. This paper briefly surveys the most usable deep learning algorithms for making the smart grid resilient, accurate, and safe. The review results show that the mentioned deep learning algorithms give excellent results over other deep learning algorithms. Therefore, these three algorithms are widely acceptable for the evaluation of smart grids.

Keywords—Smart Grid, Deep Learning, Long-Short Term Memory, Recurrent Neural Network, Convolution Neural Network.

I. INTRODUCTION

Deep learning (DL), which is based on artificial neural networks (ANN) with representation learning, has recently resurrected machine learning [1]. Because of its capabilities, different types of DL approaches are accessible in the literature, each with its own set of applications [2]. ML and DL aids in the analysis of smart grid applications in terms of efficiency and efficacy [3]. To safeguard the smart grid (SG) from cyber threats, such as collectively designing and creating appropriate SG testbeds to promote research [4]. A deep neural network is in fact an autoencoder for presenting relatively a set of data with degraded dimension [5]. Through the demand side management (DSM) system, the home energy management system efficiently schedules the appliances to achieve peak-average ration minimization, energy savings, and cost reduction [6]. To achieve the ideal set-point of smart inverters and the arrangement of capacitor banks for accomplishing dual timeframe voltage control in distribution networks, several researchers used an alternating current power flow model using deep Q networks [7].

Adversarial attacks are inherently covert and capable of causing random or targeted malicious effects by substituting artificial adversarial instances for natural inputs in a target model [8]. False data injection attack detection (FDIA) scheme based on the Kalman filter and recurrent neural network can be used for state prediction [9]. The computational cost of DL results in an unavoidable disconnect between theoretical analysis and real-time actions [10]. A graphical model-based method for detecting abnormalities in SG control systems uses Bayesian networks to map the interaction between sensors and actuators [11]. DL provides particular benefits in handling complicated issues like power system frequency analysis and control [12]. The existing ML-based FDIA detection algorithms change measured data to malicious measurement data with false data injection attacks [13]. The significance of electricity theft detection in SG is crucial for cost-effectiveness [14].

The application of DL in SG makes the power system secure and cost-effective. Even after such a great advantage of DL, it seems that SGs still have to help the power companies to enhance their operating and environmental performance. In addition, there is a research gap for the implication of DL in electric vehicles. To achieve a clean environment, electric vehicles will play an important role. If the DL concept applies in this sector, it would certainly transform the SG and helps to protect the environment. This motivates the author to do this literature review on the application of DL in SG.

II. MATERIALS AND METHODS

The WOS has been explored from the year 2011 to 2022. The inquiries include a combination of all the deep learning methods with “smart grid”. The search resulted in 1750 articles. Among them 200 had been selected as the most relevant through screening the titles and the abstracts of the manuscripts. In the second step, the classification of the method based on the deep learning method had been performed along with the PRISMA. The method of state-of-the-art review had been adapted from [15-27]. After in dept
study of the refined articles. The fundamental and novel methods of the deep learning had been identified and classified in the table 1. Further tables had been constructed based on the method of deep learning.

### III. STATE OF THE ART

The fundamental and novel methods of the deep learning had been identified and classified in the table 1.

<table>
<thead>
<tr>
<th>References</th>
<th>Year</th>
<th>Source</th>
<th>Application</th>
<th>Methods/Algorithms</th>
</tr>
</thead>
<tbody>
<tr>
<td>[28]</td>
<td>2022</td>
<td>Computers and Electrical Engineering</td>
<td>Risk management and cyberattacks detections</td>
<td>SVM, and DNN</td>
</tr>
<tr>
<td>[29]</td>
<td>2022</td>
<td>Electrical Engineering</td>
<td>Real-time exact data intrusion detection</td>
<td>State estimation, K-nearest neighbors, SCADA</td>
</tr>
<tr>
<td>[31]</td>
<td>2021</td>
<td>Computers and Electrical Engineering</td>
<td>False data cyber data</td>
<td>DNN, Agent-based model</td>
</tr>
<tr>
<td>[32]</td>
<td>2021</td>
<td>Energy Research</td>
<td>Forecasting residential energy management</td>
<td>Pooling-based deep neural network, Neural network auto aggressive integral moving average</td>
</tr>
<tr>
<td>[33]</td>
<td>2021</td>
<td>Soft Computing</td>
<td>Damage in the transmission lines</td>
<td>Convolution neural network (CNN), SVM</td>
</tr>
<tr>
<td>[34]</td>
<td>2021</td>
<td>IEEE Access</td>
<td>Effective pricing schemes</td>
<td>Reinforcement learning</td>
</tr>
<tr>
<td>[35]</td>
<td>2021</td>
<td>IEEE Access</td>
<td>Analyse accuracy, limitations</td>
<td>Federated learning, Edge intelligence, and Distributed computing</td>
</tr>
<tr>
<td>[38]</td>
<td>2020</td>
<td>Electrical Energy Systems</td>
<td>Electricity theft</td>
<td>SVM, RF, 1D-CNN</td>
</tr>
<tr>
<td>[39]</td>
<td>2020</td>
<td>IEEE Industrial Informatics</td>
<td>Uncertainty of customer’s groups</td>
<td>Bayesian deep learning; Clustering-based pooling</td>
</tr>
<tr>
<td>[40]</td>
<td>2020</td>
<td>Energies</td>
<td>Energy consumption forecasting</td>
<td>Factored conditional restricted Boltzmann machine forecasting module, Genetic Algorithm, Evolution algorithm, ANN</td>
</tr>
<tr>
<td>[41]</td>
<td>2020</td>
<td>Energy</td>
<td>Real-time economic generation dispatch</td>
<td>Expandable deep learning algorithms, Imperialist competitive algorithm, Shuffled frog leaping algorithm</td>
</tr>
</tbody>
</table>

SG is very vulnerable to cyber-attacks. This is because it is a combination of communication systems and the Internet of Things (IoT). DT and DNN classifiers can be used for constructing a unique symmetric demonstration of the asymmetric datasets of the SG control systems. This incorporates detecting attacks in SG [28]. The DL-based model in combination with the classical bad data detection algorithms can detect unstructured and structured false data inoculation attacks [29]. The SG systems manage the operation of all associated components due to the integrated communication infrastructures. The Trust-Based Iterative Energy-Efficient Routing Protocol can help in secured data transmission in the home area network [30]. An agent-based approach provides attack exposure matrices with the integration of true data. The agent-based approach also helps in the decentralization of the data integrity system [31]. The DL algorithm can also help forecast computer-aided residential energy management [32]. The DL algorithms, like CNN and SVM, can be used to identify the damages in the transmission lines [33].

The demand response modeling using the DL algorithms like robust adversarial reinforcement learning and gradient-based Nikaio-Idosa function provides an optimal strategy that exhibits the scheduling of appliances [34]. The DL has several applications in the SG paradigm. It can help to analyze the extent of limitations, awareness, prediction, feasible scenarios, accuracy, and many more [35].

In SGs, we can maximize the social welfare can be maximized by making a balance between power supply and demand through the electricity market. The increasing frequency of SG requires fast tuning strategy of neural networks to significantly handle the topology change [36]. The DL algorithms such as clustering can formulate power loss estimators for large-scale low-voltage distribution areas. These areas can be uncertainty in the smart meters, load unbalances, and power variability from distributed generation [37]. DL can be useful in the classification of non-periodicity of electricity that helps in identifying the power theft by the consumer [38]. The clustering-based pooling method can address the overfitting data gathered from the smart meters by increasing the data diversity and volume for the framework which improves the predictive performances [39]. DL-based modules improve authenticity with reasonable convergence of the predicted results [40]. DL algorithms are very useful in real-time economic generation dispatch because they provide several generation commands for future SGs with different topologies [41]. The implementation of DL-based iterative ResBlocks improves the forecasting performance [42].

#### A. Long-Short-Term Memory

Recurrent neural networks (RNN) use Long Short-Term Memory. LSTM has three gates which include the forget gate, input gate, and output gate. By using the forget gate the LSTM will find out which information needs to be forgotten and which needs to be retained by using the parameter ‘f_i’. This is based on the current input ‘x_t’, the previous output ‘h_{t-1}’, and the previous state ‘c_{t-1}. The forget parameter can be given by equation (1) [43]

$$f_t = \sigma(W_f[x_t \ h_t - 1] + b_f)$$ (1)
The LSTM helps to determine a satisfactory, and reasonable price for customers by predicting the energy demand of customers. This provides reliable service to the customers. The SG is integrated with the power grid and large-scale information and Communication Technologies and is the largest and most widely used data communication network in the IoT framework. It collects and analyzes data from distributors, transmission lines, substations, and consumer networks. The LSTM autoencoder and logistic regression classifiers identify false data injection attacks from the normal system operation events. The temporal correlations between the multi-dimensional feature vectors are used to train the LSTM-Autoencoder. [45]. A B-LSTM network is a reliable tool for time-series forecasting tasks. This handles the data with sharp variations and high stochastic behavior. The B-LSTM contains bidirectional memory-feedforward and feedback loops, which allows to see into data from previous and future hidden layers [46]. However, GA-LSTM increases the convergence speed that provides optimized effective performance, and lower execution time. This can be very useful for energy consumption prediction. When compared to random approach strategies, GA-LSTM gives better convergence. For the best answer with the least amount of error, GA constructs a new vector. [47]. The reconstruction of a new framework using the LSTM and multivariant linear regression gives better short-term load forecasting results [48]. Because of the dynamic context in which SG devices operate, the wireless link is easily disrupted, resulting in strong stochastic aspects. LSTM-NN based helps to calculate the communication link reliability confidence interval for prediction in wireless communication systems in SG [49].

Optimal prediction intervals can be constructed by using LSTM which helps in sampling data around the forecasted sample data of the SG’s components. This can help in modeling a cloud-fog architecture that can be fast, feasible, flexible, reliable, and secure for modern SG [50]. Forecasting can help participants in the electricity market compete using bidding techniques. Despite its conceptual simplicity, an LSTM-based sequence-to-sequence network can estimate power consumption and pricing for smart city time-series data with high accuracy [51]. For predicting the stability of the SG, a multidimensional LSTM algorithm can be used to make cyber-physical systems [52].

The efficiency and stability of modern SG are infeasible for power load and demand forecasting. Jaya-LSTM algorithm can be used to optimize the number of epochs, batch size, and window size. In SG, the Jaya-LSTM can be used to achieve the minimum mean absolute error. Big data is utilized to forecast electricity prices and demand using LSTM. [53]. Electricity theft has one of the severe non-technical losses. A robust CNN-LSTM model can be very useful for electricity theft detection using the synthetic minority over-sampling technique. This gives better performance in recall, accuracy, and precision. CNN is a frequently used feature extraction and classification technology that automates the process [54]. The maximum information coefficient analyzes the correlation between real-time price and load and the ARIMA-based LSTM model can predict better model accuracy than the conventional model [55].

### B. Recurrent Neural networks

The RNN is a special kind of ANN used for the evaluation of sequential data. It can generate the next output from the previous input. LSTM is a peculiar type of RNN which has better performance than others in predicting time-series data. This is due to the existence of the gate functions. The cell memory is represented in equation 2 [56]:

\[
C_t = (F_t \otimes C_{t-1}) \oplus (I_t \otimes \tanh(W_{xt} + U_{xh_{t-1}} + B_c))
\]  

(2)
Where $C$ is the cell memory, $H$ is the hidden state, $X$ is the input, $B$ is bias, $W$ and $U$ represent the weight vectors of $F$ and $H$, respectively, $tanh$ represents the tangent function.

Table III. Recurrent Neural networks in smart grids

<table>
<thead>
<tr>
<th>References</th>
<th>Year</th>
<th>Source</th>
<th>Application</th>
<th>Module/Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>[56]</td>
<td>2022</td>
<td>IEEE Internet of Things Journal</td>
<td>False Data Injection Attack Detection</td>
<td>Backpropagation module, Kalman filter, and RNN</td>
</tr>
<tr>
<td>[57]</td>
<td>2019</td>
<td>ITT 2019 - Emerging Technologies Blockchain and IoT</td>
<td>Short Term Price and Load Forecasting</td>
<td>ANN, RNN, LSTM</td>
</tr>
<tr>
<td>[58]</td>
<td>2019</td>
<td>6th, ICCSS 2019</td>
<td>Optimal energy management</td>
<td>1-layer RNN</td>
</tr>
<tr>
<td>[59]</td>
<td>2019</td>
<td>IICEI 2019</td>
<td>False data detection</td>
<td>Residual RNN</td>
</tr>
<tr>
<td>[60]</td>
<td>2019</td>
<td>Lecture Notes in Computer Science</td>
<td>False Data Injection</td>
<td>Wide and RNN</td>
</tr>
<tr>
<td>[61]</td>
<td>2018</td>
<td>IEEE, ISGT 2018</td>
<td>False data injection attacks</td>
<td>Backpropagation learning Algorithm, RNN</td>
</tr>
<tr>
<td>[63]</td>
<td>2017</td>
<td>IEEE, ISGT 2017</td>
<td>Classification</td>
<td>RNN, Hidden Markov Model</td>
</tr>
<tr>
<td>[64]</td>
<td>2016</td>
<td>Lecture Notes in Computer Science</td>
<td>Optimal real-time price</td>
<td>NN, Optimization, Tikhonov regularization item</td>
</tr>
<tr>
<td>[65]</td>
<td>2015</td>
<td>Neurocomputing</td>
<td>Optimal real-time</td>
<td>Lyapunov-like method, RNN,</td>
</tr>
</tbody>
</table>

Table IV. Convolutional Neural networks

The short-term load and price forecasting can be achieved by enhanced RNN that eliminates irrelevant features by using recursive feature elimination [57]. The battery is an important component of any SG. The absorptive ability of the battery energy storage system is different in SG as compared to the conventional energy management system. A 1-layer RNN help to solve an energy management model in SG [58]. State estimation is commonly used for the operation of SG. However, this method is not reliable during the security attacks. Yet, the residual recurrent neural network can model the anomaly and model the attack [59]. In addition, a wide and RNN model can also be used for detecting false data injection attacks [60]. The RNN-based module is superior in detecting false data injection [61]. The reliability of SG can be enhanced by detecting the anomaly in the SG. An encoder-decoder framework with RNN can detect the anomaly with an unexpected high reconstruction error [62].

To get better resilience SG it is very important to know the behavior categories of power consumers. RNN helps in the classification of consumers and makes a forecast-based classification framework [63]. Pricing is an important factor to develop an effective consumer-side management SG system. The optimized NN maximizes the aggregate utilities of all the users and minimizes the price imposed on the power provider [64]. The RNN solves the real-time price problem in the most optimized way as compared to other methods [65]. The sensor in SG holds a major role and therefore its accuracy in the system is very important. An RNN with layer feedback for each sensor provides an accuracy of the control data from different sensors. This helps to estimate the amount of fault data of the sensor by using the Kalman filter [66].

C. Convolutional Neural networks

The CNN is analogous to classic ANN as it is made up of neurons that learn to optimize. CNN has five main structures of data layer: - input to the entire NN, b) convolution layer, c) pooling layer, and d) fully connected layer [67]. In SG, the identification of faulty high-voltage power lines is very important as it leads to severe losses. These faults can happen because of several environmental hazards like severe voltage fluctuations, lighting, and incorrect design of electric field distribution. CNN and relief-F algorithms can be used to detect the power lines in SG [68]. The continuous wavelet transforms, and wavelet-CNN are used to detect the power lines in SG [69]. AlexNet-based deep convolution network helps in estimating the aging of conductor morphology of high-voltage electricity grid [70]. A CNN-based detector identifies distributed denial of service attacks on the electric vehicle charging station. Such a detector helps to model the demand values [71].

Ensemble DeepCNN can detect atypical behaviors of SG by using a random bagging method to deal with a highly imbalanced dataset. A random under bagging strategy is used to deal with the imbalance data as the first layer of the model, then deep CNNs are used on each subset, and lastly, a voting system is integrated as the final layer [72]. The measurement dataset included active power injection into IEEE39-buses and active power flow in the branches. A CNN algorithm mitigates the impacts of stealthy false data injection attacks in SGs [73]. Enhanced-CNN helps in modeling the uncertainties in time series [74]. For power price forecasting, CNN-based algorithms minimize data dimensionality and transmit crucial data to a classifier. Linear discriminant analysis reduces the data dimensionality and electricity price forecasting [75]. The CNN-based affinity propagation clustering algorithm and matching pursuit decomposition algorithms provide the cost-effective fault diagnosis in SG [76]. Electricity price and load forecasting provide future trends and consumption patterns. CNN-based algorithm and enhanced support vector regression give better results for electricity price and load forecasting [77]. Table IV shows the application of CNN in SG.
The CNN-based algorithm provides an energy management system by forecasting renewable energy resources [78]. The hybrid of CNN-based fusing XG-Boost and DT can be used to create a model for forecasting price. This constitutes in feature engineering, and classification. The enhanced CNN and support vector regression can be used in classification to evaluate the model performance [79]. The wide and deep CNN model performs best among the existing model in case of electricity theft in SG [80]. Although the [81,82] represents some advanced deep learning and machine learning methods in smart grid, a wide range of state-of-the-art methods of machine learning, e.g., [83-94], are yet to be experimented. From the hybrid methods of deep learning and machine learning to the ensemble and optimized machine learning methods, e.g., [95-109], numerous techniques have not yet been applied in the smart grid applications. For the future research using such methods for developing advanced models are suggested.

CONCLUSION

Major deep learning methods for smart grid had been represented and a new taxonomy presented. The study showed that three major deep learning method of RNN, CNN and LSTM had the most applications among the other deep learning methods. These algorithms are best suitable for risk management analysis, forecasting, electricity theft, anomaly detection, false data injection attacks and cyber security.

List of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SG</td>
<td>Smart Grid</td>
</tr>
<tr>
<td>ML</td>
<td>Machine Learning</td>
</tr>
<tr>
<td>DL</td>
<td>Deep Learning</td>
</tr>
<tr>
<td>LSTM</td>
<td>Long-Short-Term Memory</td>
</tr>
<tr>
<td>DNN</td>
<td>Deep Neural Network</td>
</tr>
<tr>
<td>RNN</td>
<td>Recurrent Neural Network</td>
</tr>
</tbody>
</table>

REFERENCES


The CNN-based algorithm provides an energy management system by forecasting renewable energy resources [78]. The hybrid of CNN-based fusing XG-Boost and DT can be used to create a model for forecasting price. This constitutes in feature engineering, and classification. The enhanced CNN and support vector regression can be used in classification to evaluate the model performance [79]. The wide and deep CNN model performs best among the existing model in case of electricity theft in SG [80]. Although the [81,82] represents some advanced deep learning and machine learning methods in smart grid, a wide range of state-of-the-art methods of machine learning, e.g., [83-94], are yet to be experimented. From the hybrid methods of deep learning and machine learning to the ensemble and optimized machine learning methods, e.g., [95-109], numerous techniques have not yet been applied in the smart grid applications. For the future research using such methods for developing advanced models are suggested.

CONCLUSION

Major deep learning methods for smart grid had been represented and a new taxonomy presented. The study showed that three major deep learning method of RNN, CNN and LSTM had the most applications among the other deep learning methods. These algorithms are best suitable for risk management analysis, forecasting, electricity theft, anomaly detection, false data injection attacks and cyber security.

List of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SG</td>
<td>Smart Grid</td>
</tr>
<tr>
<td>ML</td>
<td>Machine Learning</td>
</tr>
<tr>
<td>DL</td>
<td>Deep Learning</td>
</tr>
<tr>
<td>LSTM</td>
<td>Long-Short-Term Memory</td>
</tr>
<tr>
<td>DNN</td>
<td>Deep Neural Network</td>
</tr>
<tr>
<td>RNN</td>
<td>Recurrent Neural Network</td>
</tr>
</tbody>
</table>

REFERENCES


Dispatch in Smart Grid (2022) IEEE Transactions on Industrial Informatics.


