Engineering Applications of Artificial Intelligence 96 (2020) 104000

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Deep learning in electrical utility industry: A comprehensive review of a decade of research

Manohar Mishra [a], Janmenjoy Nayak [b,*], Bighnaraj Naik [c], Ajith Abraham [d]

[a] Institute of Technical Education and Research, Siksha 'O' Anusandhan University, Bhubaneswar, 751030, India
[b] Department of Computer Science and Engineering, Aditya Institute of Technology and Management(AITAM), Tekkali, 532201, India
[c] Veer Surendra Sai University of Technology, Government of Odisha, 768018, India
[d] Machine Intelligence Research Lab, USA

ARTICLE INFO

Keywords:
Artificial intelligence
Deep learning and machine learning electricity demand forecasting
Fault detection and classification
Power quality
Smart-grid and microgrid
Solar-photovoltaic and wind forecasting

ABSTRACT

Smart-grid (SG) is a new revolution in the electrical utility industry (EUI) over the past decade. With each moving day, some new advanced technologies are coming into the picture which forces the utility engineers to think about its application to make the electrical grid become smarter. Artificial intelligence (AI) techniques such as machine learning (ML), artificial neural network (ANN), deep learning (DL), reinforcement learning (RL), and deep-reinforcement learning (DRL) are the few examples of above-mentioned advanced technologies by which large volume of collected information being processed, and deliver the solution to the complex problems associated with EUI. In recent times, DL for artificial intelligence applications has gained huge attention in the diverse research area. The traditional ML techniques have several constrained for processing the data in raw form. However, the DL provides the options to process the raw data without extracting and selecting the feature vector. The DL techniques belong to a new era of AI development. This article presents the taxonomy of DL algorithms available in the literature applied to different problems in EUI. The main objective of this survey is to provide a comprehensive idea to the researcher/utility engineer about the applications and future research scope of DL methods for power systems studies.

1. Introduction

The electrical utility industry (EUI) is a very complex artificial system. The EUI has involved constantly in exploring ways to upgrade the efficacy and dependability by means of which it delivers energy. Even if the elemental technologies of energy generation, transmission, and distribution revolutionize very gradually, the EUI has been hurried to explore novel tools that may possibly help to show the benefits (Warwick et al., 1997; Madan and Bollinger, 1997). The EUI is a sector that ensures a consistent production and supply access of electricity to consumers. To achieve this, the utility companies are also adapted a trendy strategy called as Digitalization. In digitalization process, several modifications have been employed to yield the capability to generate new path of EUI for value creation. The fourth industrial revolution has also begun to provide a shape that enables the decentralized intelligence in manufacturing and production. The electrical power system is in the focus of a revolution, as technology and modernization disturb conventional models from generation to afar the matter. The global electrical power sector is in a process of changing its traditional technologies for the generation, transmission and distribution of electrical power through integrating new digitally-advanced and green technologies. In this regards, the following are the major three trends in particular to achieving the game-changing revolutions: (i) Electrification (ii) Decentralization (iii) Digitalization. The main goal of electrification tend is to achieve a long-term carbon goal through triggering the relevant distributed resources and electrifying the large sector of economy through heating and transport. This can be realized through the technologies such as electric vehicle, vehicle to home/grid, smart charging etc. The decentralization trend leads to a changeover to a more digital and interconnected power system that will handle the central generation in concert with distributed energy resources (DERs). Once the industry seems to explore possibilities outside the conventional electrical-grid, DERs like wind energy system (WES), solar photovoltaic system (SPV), and hi-tech energy storage system, it can makeover an old-fashioned power-grid into a smarter, interconnected system named as smart-grid (SG). Similarly, the digitalization of electrical energy system allows for open, real-time automated communication and automation of the system (Angelopoulos et al., 2019).

The main objective of EUI can be summarized as follows: optimum resource utilization, higher energy efficiency, higher system reliability, higher system security and economical electricity distribution to

* Corresponding author.
E-mail addresses: manohar2006mishra@gmail.com (M. Mishra), mailforjnayak@gmail.com (J. Nayak), mailtobnaik@gmail.com (B. Naik), ajith.abraham@ieee.org (A. Abraham).

https://doi.org/10.1016/j.engappai.2020.104000
Received 24 December 2019; Received in revised form 19 August 2020; Accepted 4 October 2020
Available online 9 October 2020
0952-1976/

consumers. The traditional methods for electrical power system (EPS) analysis, control and decision making are mostly supported by physical modelling and numerical calculations. But, this procedure usually faces difficulty in addressing uncertainty and partial observability problems, and therefore, they cannot come across the necessities of future growth of SGs. In the area of EPS, one of the highly rousing and possibly cost-effective latest advances is the increasing use of artificial intelligence (AI) techniques. The AI techniques, such as expert systems (ESs), machine learning (ML), fuzzy logic (FL), artificial neural networks (ANNs), and reinforcement learning (RL) are the few examples of cutting-edge technologies by which large volume of collected information being processed, and deliver the solution to the complex problems associated with SG. These technologies support effective tools for design, simulation, control, estimation, fault diagnostics, and fault-tolerant control in modern EPS (Bose, 2017).

Under the umbrella of AI, deep learning (DL) and deep-reinforcement learning (DRL) are the newest technology in the last few years. Recently, deep learning (DL) concept in artificial intelligence (AI) application has gained huge attention. The DL/DRL techniques belong to a new era of AI development (Zhang et al., 2018). DL states about the architectures which comprise numerous hidden layers (deep networks) to learn several distinct features with multiple levels of perception. The traditional ML techniques have several constrained for processing the data in raw form. However, the DL provides the options to process the raw data without extracting and selecting the feature vector (Wani et al., 2020). Fig. 1shows this concept in pictorial form. In place of handcrafting numerous rules and procedures for the extraction of feature vector from raw natural data, DL involves learning these features automatically at the time of the training period (Wani et al., 2020; LeCun et al., 2015; Goodfellow et al., 2016). Hence, DL, RL, and DRL seem to be some of the facilitating tools for the forthcoming progress and achievement of EUI.

The manuscript offers an organized and comprehensive overview of the DL study on the electrical utility industry (EUI), which covers several application fields. Previously published articles on this subject either emphasized on a specific application domain or a particular research field. Most of the accessible works that have published earlier focuses its efforts on image processing applications; there is an increased concern from the electro-mechanical domain. Therefore, this manuscript provides a comprehensive literature review on developments of DL techniques and its variant's, and application of DL in several electrical domains (such as, power system fault detection and classification, load and power forecasting, wind speed and irradiance forecasting for wind and PV system respectively, power quality detection, power system state estimation, etc…). The previously reported research articles have been distributed in many databases. We have chosen some standard electronic databanks to generate a comprehensive bibliography of a research paper on DL architectures in EUI. The objective of this article has been fulfilled by reviewing several journal articles and conference papers, which are directly related to the DL applications in the electrical power domain. The major goal of this study is to recognize the recent research of DL in EUI. This is achieved by means of exploring obtainable printed articles that provide understandings of prospective applications and researcher concerns on the foremost leanings, substantial works, and future research scope. Thus, we have tried to gather an organized reference opinion for booming literature of this emergent research area. As the studies within this topic have practical significance, the scope of this search includes the year span of 2010–2020. This objective of this article has been fulfilled by reviewing several journal and conference articles those are straightforward related to the DL applications in the electrical power domain. With this regard, the literature search was accomplished via electronic databases such as Science Direct, IEEE Xplorer, Scopus, IET Digital Library, Springer, Willy, etc... The principal descriptor applied is "deep learning", clustered with the following: "electrical power system", "electrical utility", "electrical fault identification", "power quality", and "microgrid", "energy management". Over-All, the authors have recognized 115 articles that were published especially on DL application to the electrical utility industry. A remarkable observation is the lack of survey articles on this particular area. Therefore, we have considered this as a superior opportunity to help the researchers working on this particular domain through imparting a comprehensive up-to-date survey article. This article may act as a reference point of the relevant future research carried out by the reader.

With an intention to provide a comprehensive and most relevant review a proper article selection approach (ASA) is highly necessary. Therefore, we have designed an ASA for this particular review, where several inclusions and exclusion criteria are considered for designing the ASA which as follows:

The inclusion criteria for selecting the most appropriate publications as:

1. Addressing the types of deep learning architectures applied in EUI.
2. Addressing the DL architectures in electrical research domains
3. Addressing the major impact and open research issues of DL application in EUI.

The exclusion criteria for removing unwanted publications as follows:

1. Articles published in books, Ph.D. and/or Masters' dissertations, meta-analysis, and other types of literature reviews.
2. Articles those are not related to EUI and concentrated on DL application for different domains such as agricultural engineering, aquaculture, transportation technology, traffic management, smart cities, home automation, computer vision, healthcare management, and medical image classification.
3. Abstracts or full manuscripts are not accessible.

In this work, a total of 2378 technical research articles have been collected to synthesize a comprehensive review article on DL application to EUI. In the initial stage, all the duplicate articles were removed from the database, which yields 522 potentially relevant articles. Subsequently, the titles and abstracts of each article were evaluated and most relevant articles were differentiated based on our objective. The outcome of this stage provides 254 articles. Lastly, the full-text reviews of the residual documents were carried out and 115 articles were selected regarding DL application in the electrical domain.

Moreover, the article discusses the computational complexity of each DL technique, taking note of their strong points and weakness.

Lastly, it is valuable to draw out the objective and significant contributions from the article:

- This manuscript presents a general idea for the academia/research society of the several architectures along with allied implementation challenges.
- This article offers an organized review of increasing literature on the application of DL in electrical domains.
- A large spared DL research scope or application area within electrical power engineering domain such as power system fault detection and classification, load and power forecasting, renewable power generation forecasting, power quality detection, power system state estimation, etc… are considered in this study.
- The main objective of this survey is to provide comprehensive idea to the researcher/utility engineer about the applications and future research scope of DL methods for power systems studies.

The rest parts of the articles are organized as follows: Section 2 describes a brief taxonomy of Deep Learning Techniques. Section 3 presents the taxonomy of the DL application domain corresponding to electrical engineering. Section 4 reviews the DL application to advanced forecasting problems. Section 5 reviews the literature of DL application in automatic power quality monitoring. Sections 6 and 7 studies the DL application feasibility in Microgrid and Electric vehicle area respectively. In Section 8, several miscellaneous areas under the electrical domain are analysed regarding the DL application. Section 9 discussed the outcomes from this survey, challenges and future scope. Section 10 concludes this survey with several concluding remarks.

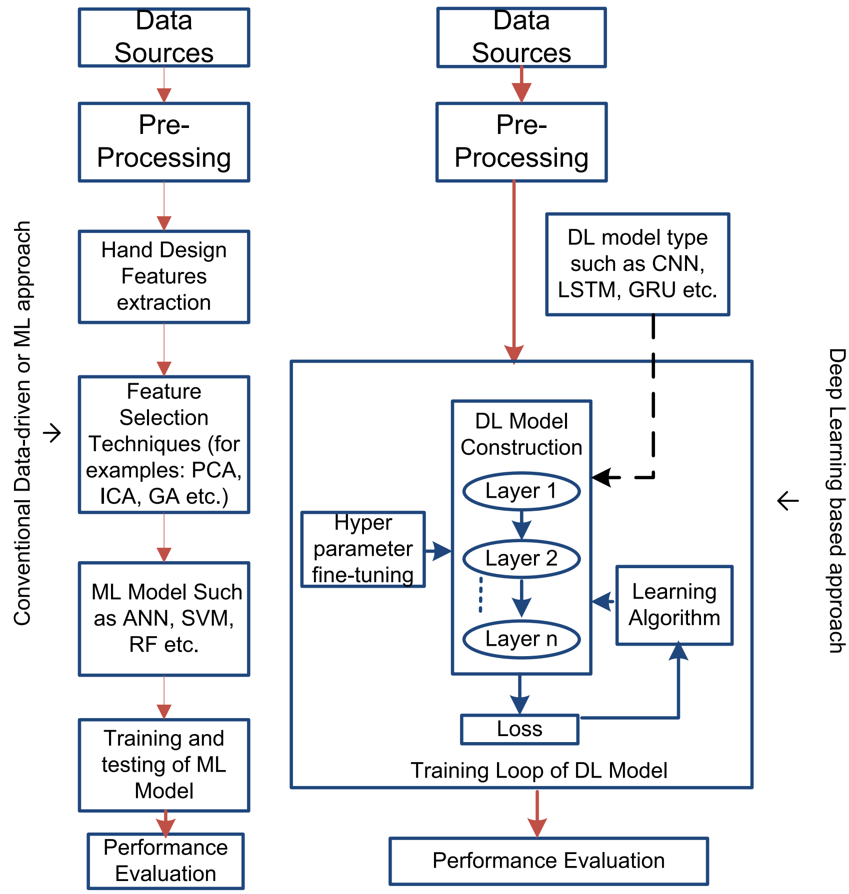


Fig. 1. Conventional ML vs. DL.

2. Taxonomy of deep learning techniques

Deep learning techniques are belonging to a new era of artificial intelligence (AI-2.0) (Zhang et al., 2018). Under the umbrella of AI, DL is the newest approach in the ML framework (Mishra et al., 2020). Deep learning applies either deep or hierarchical learning approaches, which is a class of ML renovated mostly from 2006 onwards. In recent times, the DL concept in AI applications' has gained huge attention. In contrast to conventional ML, the DL uses multiple layers for the extraction of a higher-level feature vector progressively from the raw input data. Fig. 1 portrays the general conceptual difference between conventional ML and DL. Considering the advantages DL architecture over ML, the same has been successfully used in diverse fields of applications such as speech processing (Wang et al., 2016a), image processing (Wu et al., 2018), natural language processing (Deng and Liu, 2018), visual art processing (Xie et al., 2017), health care management (Khan and Yairi, 2018), military (Dijk et al., 2019), etc.

DL methods can be classified (Alom et al., 2019) as: (i) supervised, (ii) semi-supervised, and (iii) unsupervised. Furthermore, there is one more class of DL approach named Reinforcement Learning (RL) or Deep reinforcement learning (DRL). Fig. 2 presents the pictographic representation of taxonomy of DL architecture available in the literature.

2.1. Deep Supervised Learning (DSL)

The DSL technique generally uses labelled data in their algorithm. In this type of learning process, the algorithm handles a set of inputs and corresponding outputs $(x_t, y_t) \sim \rho$. Let us consider an example where the input to the algorithm is 'x_t', prediction output is $\hat{y}_t = f(x_t)$, and the subsequent loss value is $l(y_t, \hat{y}_t)$. Afterwards, intelligent agent of the DSL will adjust the network parameters for an improved approximation of the required outputs, and this has been achieved through several iterations. After successful training, the algorithm will be able to get the correct prediction with less error. In DSL framework; Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), including Long Short Term Memory (LSTM), and Gated Recurrent Units (GRU) are the different DL techniques available in literature. The detail mathematical analysis about all these DL networks can be extracted from several previously published articles such as DNN (Miikkulainen et al., 2019; Bouwmans et al., 2019), CNN (Yao et al., 2019; Zhou, 2019), RNN (Ding et al., 2018; López et al., 2016), LSTM (López et al., 2016; Wang and Li, 2018a) and GRU (Wani et al., 2020; LeCun et al., 2015; Goodfellow et al., 2016).

2.2. Deep Semi-supervised Learning (Deep-SSL)

The SSL process has been designed for partially labelled data. In a few instances, deep reinforcement learning (DRL) and Generative Adversarial Networks (GAN) have been used as Deep-SSL methods. The mathematical analysis of GAN was detailed in Mao et al. (2017). Section 2.4 provides a summary of DRL methods. Moreover, RNN, including LSTM and GRU, can also be used as Deep-SSL techniques.

2.3. Deep Unsupervised Learning (Deep-USL)

The Deep-USL systems are such types of learning that do not dependent on data labels. In this type of learning approach, the agent learns the internal characteristics or key features to determine unknown interactions or structure within the input datasets. In the case of clustering approach, the dimensionality reduction, and generative approaches are contemplated as USL methods. There are numerous agents under the

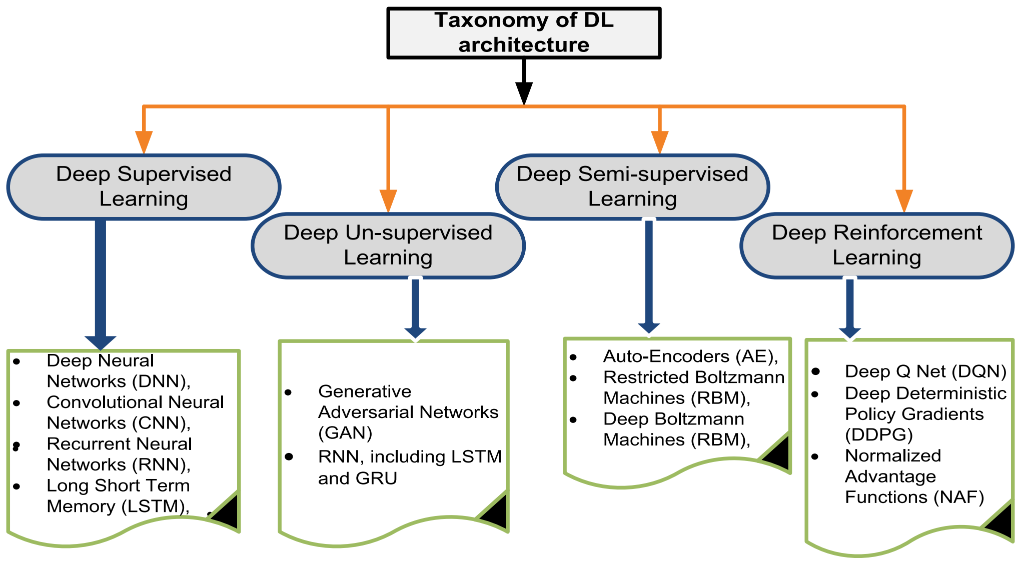


Fig. 2. Taxonomy of DL architecture available in literature.

umbrella of DL framework that are worthy at clustering and non-linear dimensionality reduction, such as Auto-Encoders (AE), Restricted Boltzmann Machines (RBM), and GAN. The mathematical analysis of AE and its' variants, and BM and its' variants were detailed in Wani et al. (2020), LeCun et al. (2015) and Goodfellow et al. (2016). Furthermore, RNN networks, for example LSTM and reinforcement learning may also consider in different USL application fields.

2.4. Deep Reinforcement Learning (DRL)

RL is a subset of ML supported with sequential decision making under uncertainty. The major objective of RL is for the agent to maximize the cumulative reward by taking a series of actions in response to a vibrant environment. Like supervised and unsupervised learning, the DL is also a basic ML model. The Q-learning, SARSA (State–Action–Reward–State–Action), DQN (Deep Q Net) and DDPG (Deep Deterministic Policy Gradients) are few examples of RL approaches. In addition, the RL approach combined with the perception of DL has led to an advanced technique called as DRL. It is mostly used for unknown environments, and can be implemented for a variety of tasks involving both rich perception of high dimensional raw inputs and strategy control. The initial development of DRL was initiated in 2013 with Google Deep Mind. Thenceforth, numerous innovative approaches have been suggested using RL. Let us consider an example of reinforcement learning: If environment input samples: $x_t \sim \rho$, agent forecast: $\hat{y}_t = f\left(x_t\right)$, agent receive cost: $c_t \sim P\left(c_t \left| x_t, y_t\right.\right)$, the environment requests to an agents for the response of a question, and based on the received answer a noisy score is allotted. This approach can be stated as semi-supervised learning in few cases. Considering diversified scope and complexity of problem, the suitability between RL and DRL has been chosen to solve the particular problem. For example, if the problem needs to optimize a massive number of parameters then DRL will be the effective one. The detail mathematical analysis of RL and DRL techniques can be from Yu and He (2019), Xu et al. (2019) and Botvinick et al. (2019). Table 1 provides a brief survey of the NNs shaping the DL architectures applied to the electrical utility industry in the literature.

3. DL for electrical power domain

Several research articles/magazine/dissertations have been published, comprising deep learning, reinforcement learning and deep-reinforcement learning applications in electrical engineering domains. The majority of publications were indexed by journals/conferences issues since 2014. The application area or necessary functional/services include advanced power system monitoring and diagnostics, adoptive protection schemes, distributed power system management, islanding possibilities in microgrids, advanced forecasting support, electric vehicle, power quality monitoring and almost all the technical fields of EUI. Fig. 3 shows a pictorial representation of the taxonomy of the application area where the DL has successfully applied till date. The details reviews related to DL application on each mentioned area are presented comprehensively in the subsequent sections.

Table 1 summarizes several application fields stated by authors of previously published papers related to different deep learning architectures in electrical domain.

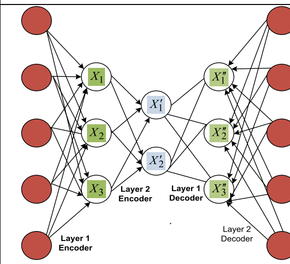
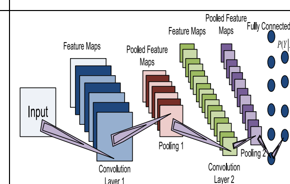
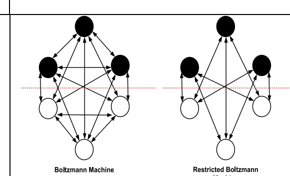
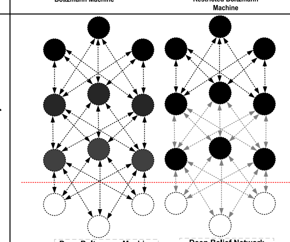
4. Advanced forecasting support

Advanced forecasting support is a vital feature for the EUI and highly essential for optimized grid operation. Several sub-areas within the electrical domain related to forecasting provisions are includes (i) load/demand prediction (ii) Power system state forecasting (iii) renewable energy generation forecasting such as wind speed, wind power, and solar irradiance, etc. The importance of AI, implementation challenges and solutions for all aforementioned application domains are already reviewed by the researchers in the last couple of decades, such as load/demand prediction in Baliyan et al. (2015), Metaxiotis et al. (2003) and Raza and Khosravi (2015), Power system state forecasting in Vankayala and Rao (1993) and Kumar and Srivastava (1999) and renewable energy generation forecasting in Kalogirou (2001), Lei et al. (2009), Elsheikh et al. (2019) and Das et al. (2018). Although the scope of DL application in these domains is not fully explored yet, the authors in this article have tried to frame the outcomes and challenges related to the research work that are already published in the same domain using deep learning techniques.

4.1. Load/demand prediction

Although the perception of SG is to upgrade the existing grid towards a more robust, reliable and well-organized grid, reducing the production cost is also an important feature. This can be achieved by proper planning and operating of EUI. Moreover, the high perpetration of DERs into the existing grid also increases the uncertainty and challenges to SG operations and scheduling. Hence, the correct

Table 1
Brief literature review on neural networks shaping the DL architectures applied to the electrical utility industry (Mnih et al., 2015; Yoo et al., 2014).

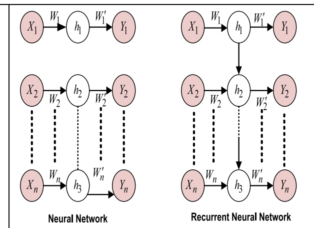
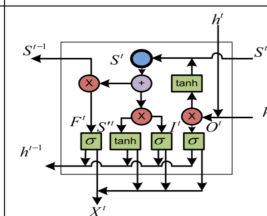
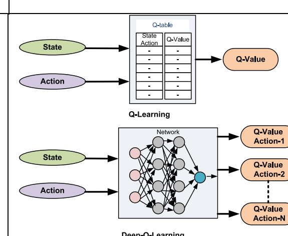
DL model	Brief Descriptions	Strength	Limitation	Structure of the Model
Auto Encoder (AE)	AE is an unsupervised learning process. It uses backpropagation for setting the target values to be equal to the inputs $\{y^{(i)}=x^{(i)}\}$. The structure of AE is comprised of a decoder and a coder, where the earlier one is used for transforming the input to an underlying representation, and the later one is for the reconstruction of input from the decoder output representation. AEs are trained to reduce reconstruction errors. By compelling the dimension of the latent representation to be different from inputs (and subsequently from the output), it is feasible to learn relevant patterns in the dataset. AEs are categorised by De-noising AE (DAE), Sparse-AE (SAE), Variational-AE (VAE), Contractive-AE (CAE).The AE, DAE, and SAE can be stacked to make deeper forecasters, which are named stacked autoencoder (SAE), stacked de-noising autoencoder (SDAE) and stacked sparse auto-encoder (SSAE), correspondingly.	• It can be amended to learn more affluent representations. • Implementation is easy. • Dimensionality reduction. • Simpler to trace the loss/cost function that is being brush off by back-propagation.	• The training process of autoencoder requires lots of dataset, high processing time and fine tuning • It usually acquires more information compared to important information. • It may perhaps not able to define important and relevant information.	
Convolution Neural Network (CNN)	CNN, like NNs, is made up of neurons with learnable weight and bias. But, unlike NN, the input to the CNN is a multi-channel image. The simple CNN structural design comprises of one convolutional and pooling layer, optionally followed by a fully-connected layer for supervised prediction. Practically, CNNs are comprised of more than ten numbers of convolutional and pooling layers to represent as an improved learning model. It is mostly applied for computer vision and image processing domain.	• It shows better performance in multi-dimensional dataset • It is very good in extracting most suitable features.	• It has more complicated architecture, and therefore provide slower training rate.	
Restricted Boltzmann Machine (RBM)	RBMs are the generative stochastic models that learn a probability distribution over the input space (Yoo et al., 2014). RBM is a variant of BMs, beside the constraint that its neurons must forma bipartite graph.	• BMs have two groups/parts such as visible and hidden. In contract to BM, the RBM have no connection between nodes within a group. It helps for more efficient training compared to general BM.	NA	
Deep Boltzmann Machine (DBM) and Deep Belief Network (DBN)	The DBM and DBN are two variants of RBM. DL structures achieved by stacking RBMs are named as DBMs.	**DBM:** • It can learn remarkable generative models. • It keeps ample required data found in deep belief network. • Here, parametric optimization can possible for all layers simultaneously. It may includes uncertainty about equivocal inputs **DBN:** • It is most suitable for 1-D data. • It is able to extract large-scale features from the input data. • It can firmly attain high performance on handling complex data without data preparation step. • It can be more effective compared to PCA in separating the more significant data.	**DBM:** • The training process is slower compared to DBN. • Slower speed of DBM limits their performance and functionality. • Approximate inference is slower reducing the application of employing DBM's for extracting feature representations. • This also makes the joint optimization impractical for large dataset. **DBN:** • Training time involved with DBN is very high ineffective	

(continued on next page)

prediction of energy/load demands at different levels is highly essential for the EUI in economic aspects. The accurate load forecasting generally depends on the knowledge of influencing factors for increasing/decreasing demands. Authors in Khatoon and Singh (2014) studied several influencing factors on electrical load forecasting. These factors are classified as: Meteorological Factors (Climate, weather, temperature, humidity, solar radiation etc.), Temporal or Calendar factors (Hours of day, days of week, and timings of year etc.), Economy Factors (Industrial development, GDP, etc.), Random Factors (Activities, Festival etc.), Customer Factors (Type of consumption, Size of building, Electric appliances, Number of employees etc.) etc. Cavallaro (2005). Authors in Xue and Geng (2012) categorized the influencing factors on electrical load forecasting into three categories, such as short-term influence factors, middle-term influence factors, and long-term influence factors.

Considering these time horizon factors, the types of load forecasting can be divided into three categories, namely Short-term forecasting

Table 1 (*continued*).

Recurrent Neural Network (RNN)	The RNN structure is designed to overwhelm the limits of simple NN by utilizing past input information to get the output. As, a classical NN architecture does not allow for the sequential dataset such as $X_1, X_2 \cdots X_N$; their corresponding hidden states $h_1, h_2 \cdots h_N$ are independent of each other. However, in an RNN structure, the hidden state at each step depends on the hidden state of the previous one. Several variations of RNNs are accessible from literatures such as LSTM, GRU... etc. These variants were presented to address the vanishing or exploding gradient problem in RNN.	• Suitable for sequential dataset • It can detect changes over time	It may face difficulty during training and real-time implementation	
Long Short Term Memory Network (LSTM)	The LSTM network is a variant of RNN. A common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate.	• It prevents back-propagated errors from vanishing or blowing up. • It can avoids the "vanishing gradient problem" • It can be operated for signals comprising both low and high frequency components. • It can learn to identify context-sensitive languages compared to other similar type's methods.	• RNN and LSTM are difficult to train as they necessitate memory-bandwidth-bound computation, which is the most horrible situation for hardware scientist and eventually bounds the application ability. • It requires more memory to train • It is easy to overfit • It is sensitive to different random weight initializations	 Where, S'^{t} (input modulation gate), I^t (input gate) and O^t (output gate) of LSTM network
Deep Q-Learning (DQL)	DQL is a variant of DRL. ***Q*-learning (QL)** is a model-free RL algorithm (Mnih et al., 2015). The objective of QL is to learn a policy, which states an agent what action to take under what environments. Applying DL architecture with RL algorithms like QL, a dominant model (DRL) i.e. DQL can be formed which is able to clear the earlier insoluble issues. The fundamental step of DQL is that the initial state is fed into the NN and it yields the Q-value of all likely actions as on output.	DRL involves building a DL model which enables function approximation between the input features and future discounted rewards values also called Q values.	• DQL is expensive for the agent, particularly, in the commencement stage. Because, every state-action pair should be visited frequently in order to converge to the optimal policy. • In comparison to DL, target variable is not stationary and therefore, training is unstable.	

(STLF), Mid-term forecasting (MTLF) and Long-term forecasting (LTLF) (Almalaq and Edwards, 2017; Haque and Kashtiban, 2000). Fig. 4 shows the details about the objective and time spans involved in these types of time horizon forecasting approaches. Several methods were already reported earlier for electricity load forecasting using AI-1.0 that is ML algorithms. The ML models such as different variants' of ANN and SVM were widely used for load forecasting in a couple of decades (Khuntia et al., 2016; Metaxiotis et al., 2003; Daut et al., 2017). These methods were based on small datasets without considering the larger dimensions of dataset extracted by smart meters (SMs). Moreover, the ML techniques may be treated as inefficient learning models for larger volumes of the dataset (Wani et al., 2020). In contrast to this DL, algorithms are highly efficient for handling high dimensional data (Wani et al., 2020). Therefore, this template presents a synopsis on the DL based methods used in SG load forecasting (Almalaq and Zhang, 2020).

4.1.1. Short-term load forecasting using DL overview

The main objective of accurate STLF is mentioned in Fig. 4 with forecasting time-period. Generation scheduling can be achieved by the STLF to determine the generation resource allocations, working limits, environmental and apparatus handling constraints, and optimal operational state of EPS. The STLF is used for ensuring the EPS security, stability and reliability. In this regards, although several work have already reported for STLF problem using conventional ML techniques, the popularity of DL application have increased in few years.

Authors in Guo et al. (2018) presented a DNN based approach for STLF. Here, the data or patterns of electricity consumption were extracted with respect to temperatures, months of the year and days of the week. Afterwards, the probability density of load consumption was forecasted by means of DNN combined with quantile regression (QR). Lastly, the outcome of the suggested approach was compared with popular ML tools such as random forest (RF) and gradient boosting machine. Authors in Hossen et al. (2017) suggested a DNN based methodology for day-ahead electricity consumption prediction. Ninety days of Iberian utility market data was used for the training of multi-layer DNN. A number of activation functions set (comprising different types of activation functions) were tested to obtain an improved Mean Absolute percentage error (MAPE) considering the weekday and week-end variants. The outcome of the study shows that the grouping of an Exponential linear unit (ELU) with ELU has better performances than other combinations when assessed via MAPE standards for weekday data-sets; while for weekend data-sets, the ReLU–ReLU (Rectifier linear unit) combination outperforms the other combinations. In Din and Marnerides (2017), Authors have studied the application feasibility and compares the prediction accuracy of the Feed-forward DNN and Recurrent-DNN models utilizing time-variants STLF data.

Researchers in Dong et al. (2017) used the combination of CNN and K-Means algorithms to predict hourly load demands. The K-means

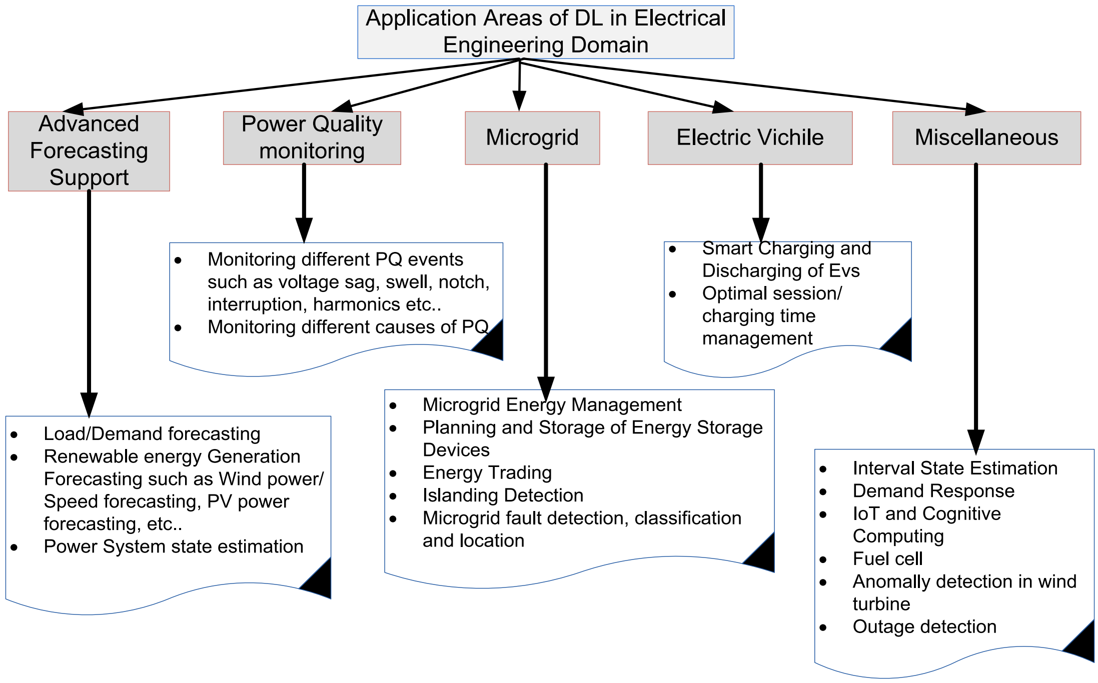


Fig. 3. Different research areas of deep learning application under the umbrella of electrical engineering domain.

algorithm was applied to culture bulky data-sets (which comprise above 1.4 million of electricity loads records) into subsets. Afterwards, these subsets were used as input to CNN for training and testing purposes. The experimental outcome of the suggested approach has clearly shown its effectiveness. Author in He (2017) proposed novel DNN architectures to predict one day-ahead hourly electricity loads. The authors use multiple CNN modules to learn deep features from the historical load dataset. Here, the max-pooling was carried out with size 2 and stride 2 and the combined outputs in one array fed into the LSTM based RNN module. Correspondingly, some other features like temperature and holidays are projected into vector representation using dense components. Lastly, all the features concatenated via dense layers to guess load amount. The stated method was evaluated using the three years of hourly data in a north China city and proven its effectiveness through experimental results.

Authors in He et al. (2017) proposed DBN combined with parametric Copula models to predict the load demand in an hourly basis in an electrical grid. A couple of forecasting scenarios such as 1-day in advance and 1-week in advance were conducted employing the stated approach. One year of electricity load consumption data in an urban area in Texas, United State was considered for the experimental validation. Experimental results, as well as a comparative analysis with other contemporary techniques (such as ANN, SVR, and ELM), proved that the approach outperforms by means of MAPE and RSME. Researchers in (Dedinec et al., 2016) suggested a DBN based method for electric load forecasting. Here, the DBN was simulated from multiple layers of restricted Boltzmann machines (RMBs) and the layer-by-layer unsupervised training process was avid by fine-tuning of the parameters by using a supervised back-propagation training method. The validation of the stated approach was successfully tested by Macedonian hourly electricity consumption data in the period 2008–2014.

Authors in Wen et al. (2020) presented an approach using DRNN–GRU models for STLF and MTLF. The France metropolitan's electricity consumption data were used for testing and validation of the stated approach. The work was accomplished by training numerous linear and non-linear ML algorithms and selecting the best as a reference point, opting the finest features sourcing wrapper and embedded feature selection techniques and lastly using a genetic algorithm (GA) to find optimal time lags and a number of layers for LSTM model predictive performance optimization. Authors (Bedi and Toshniwal, 2018) have proposed empirical mode decomposition (EMD) based DL method which combines the EMD technique with the LSTM architecture model to predict the electric load consumptions demand for a definite time interval such as day-ahead, hour-ahead, etc. In this work, the EMD process decomposes the load data basically a time-series signal into a number of intrinsic mode functions (IMFs) and residue. Afterwards, each IMF is used as input for training the LSTM separately. At last, the prediction outcomes of all IMFs are pooled together to decide a combined output for electricity demand. The EMD method combined with DBN comprising two RBMs was used to predict the load demand of Australian Energy Market in Qiu et al. (2017) by Qiu et al. Here the effectiveness of the proposed approach was proven by comparing the obtained results with other nine existing forecasting methods such as Persistence, SVR, ANN, DBN, RF, EDBN, EMD–SVR, EMD-SLFN, and EMD-RF. Estebsari and Rajabi (2020) have proposed a single residential Load Forecasting using DL especially CNN. Here, the authors have used image encoding technique integrated with CNN to decrease the mean absolute percentage error up to 40%. They have also compared their technique with other baseline approaches like SVM and ANN.

Authors (Tong et al., 2018) have proposed a DL based model for day-ahead electricity forecasting. In the initial stage, the historical load data, as well as corresponding temperature data, were refined by stacked denoising auto-encoders (SDAs). In the next stage, the output of SDAs data was used as input for the training process of support vector regression (SVR) model. The validation of the stated model was carried out by a comparative performance analysis with plain SVR and ANNs models.

The STLF at the individual building (residential customers' levels) can considerably support the smooth operations of EUI. Including accurate load forecasting strategies, the peak load shaving can be accomplished through co-operatively employing energy storage systems (ESS) or smart demand response technologies. From the EUI's viewpoint, if the exact load predictions at the residential customer's level are accessible, then the utility suppliers can target the appropriate groups of customers based on the extracted forecasting information for the active participation in demand response scheduling in the actions of energy shortage. In the meantime, SMs are available in the market

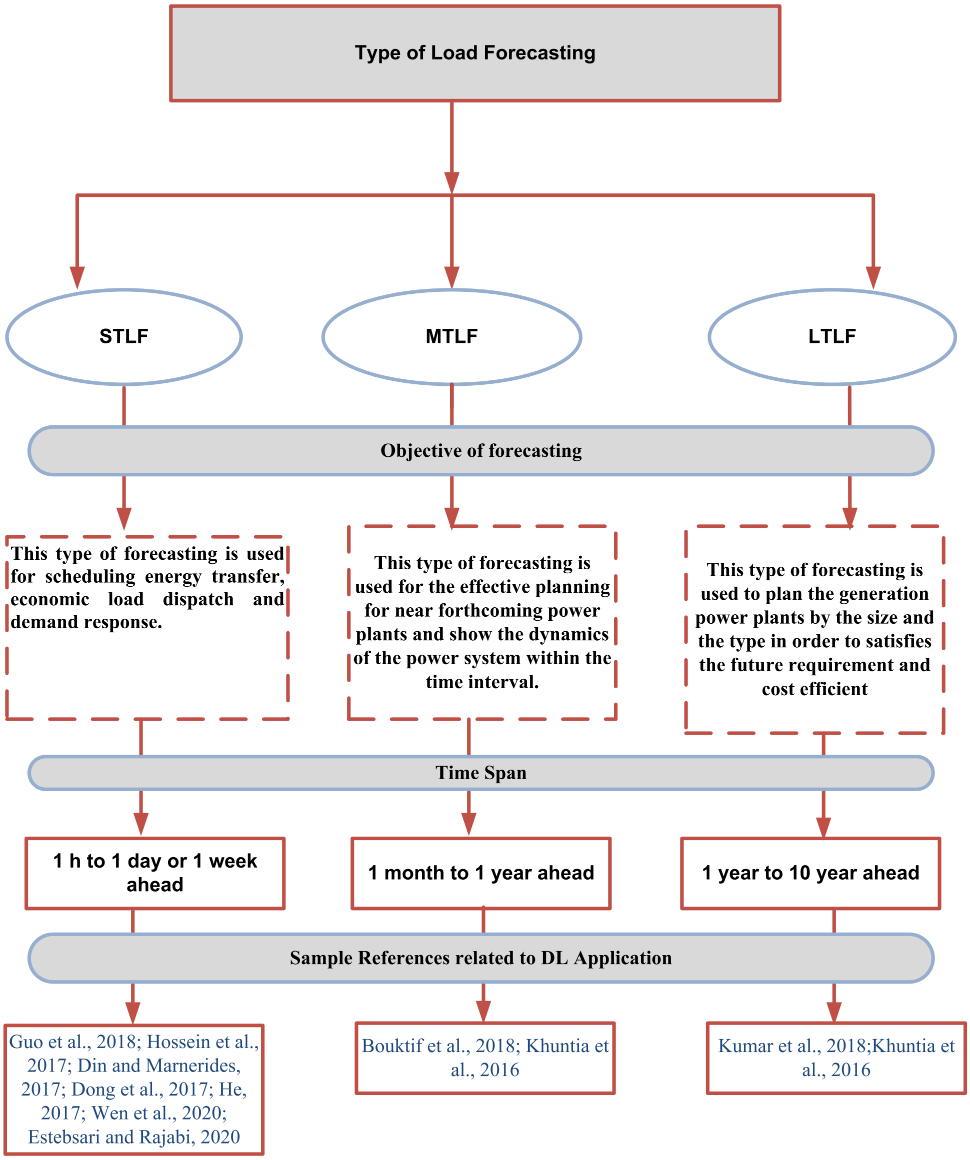


Fig. 4. Taxonomy of Load and demand forecasting methods based on time horizon.

for the collection of large energy consumption datasets at residential customer's levels, which facilitate DL based load/demand forecasting. Authors in Marino et al. (2016) proposed two LSTM architecture based hourly ahead and minute-ahead load demand prediction. Here, the result analysis proves that a standard LSTM architecture cannot accurately predict a one-minute load while the S2S LSTM-based architecture attains outstanding prediction in each case. Similarly, in Amarasinghe et al. (2017) similar authors used CNN for load demand predictions at residential building levels using historical data. Prediction outputs form the CNN are compared with those obtained by S2S based LSTM, FCRBM, "shallow" ANN and SVM for the same dataset and it shows that CNN outperforms SVR while producing comparable results to the ANN and DL methodologies.

Kong et al. addressed the STLF problem at the individual building level (Kong et al., 2019). Initially, the electrical load demand on such granular level and substation levels were compared. Afterwards, a density-based clustering method was used to calculate and compare the inconsistency between the combined load and individual loads. Generally, the lifestyle of the utility customers decides the energy consumption pattern, which is very inconsistent in nature. For that reason, the author proposed an LSTM–RNN based load forecasting structure for this type of extremely challenging load demand dataset (Kong et al., 2019). The prediction outcomes clearly prove its effectiveness by showing better results compared to other contemporary methods applied to similar datasets.

Authors in Mnih et al. (2012), Taylor et al. (2011) and Mocanu et al. (2016a) used the Conditional-RBM (CRBM) and Factored-CRBM (FCBRM) techniques to forecast residential levels loads/demand.

Shi et al. have proposed an innovative pooling-based deep-RNN for household load forecasting (Shi et al., 2017). In essence, the proposed work addressed the over-fitting issue caused by increasing data diversity and dimensions. The suggested STLF model was employed on Tensor-flow deep learning boards and tested on 920 Ireland residential datasets extracted by SMs. Compared with the other related algorithms in residential load/demand forecasting, the suggested approach outdoes ARIMA by 19.5%, SVR by 13.1% and classical deep RNN by 6.5% in terms of RMSE.

4.1.2. Mid-term and long-term load forecasting using DL overview

The main objective of MTLF and LTLF with the time span is presented in Fig. 4. The main goal of MTLF is to maintain the balance between demand and generation by proper maintenance scheduling, coordination of load dispatch and price settlement. Similarly, the objective of LTLF is to expand the EPS units.

From the 60s of 19th century, most of the load forecasting methods established till-date is committed to STLF, and not much for MTLF/LTLF. The MTLF and LTLF are much more complex than easily fitting a mathematical model to certain data, and it necessitates a great deal of knowledge about the 'substantive' problem. In comparison to STLF that usages several workout on data modelling (for instance, fitting prototypes to data-sets and concluding from them, deprived of additional knowledge on the working of EPS), the MTLF and LTLF are hinged on enriched expertise on modelling, and working skill with EPS, and an in-depth insight of EPS working, and in what way the electrical electricity industry may get affected by the fluctuations in a national economy all over the years, or by modifications in technology and so on.

Therefore, the current availability of research work based on MTLF and LTLF as compared with STLF is hard to notices because of uncertainty, complexity and difficulty in collecting as well as processing of dataset. Authors (Khuntia et al., 2016) have presented a detail analysis of these issues with few available approaches to forecast mid- and long-term load demand. The paper analysed the forecasting approaches based on parametric and non-parametric (or AI based). Linear regression, autoregressive integrated moving average (ARIMA), and grey dynamic models are few examples of parametric forecasting approaches that reported for MTLF and LTLF in past (Abdel-Aal and Al-Garni, 1997; Barakat, 2001; Filik et al., 2011; Kang and Zhao, 2012). Similarly, the non-parametric or AI based approaches like fuzzy logic clustering neural network, fuzzy linear regression model, support vector regression, optimization algorithm based model etc. are reported for MTLF and LTLF in You et al. (2006), Yue et al. (2007), Wang et al. (2012) and Lee et al. (1997).

In the meantime, the current applications of DL based algorithm in MTLF and LTLF problem have reported in few articles as follows. Authors in Bouktif et al. (2018) presented an approach using ML and LSTM models for STLF and MTLF. The validation of the suggested approach was tested with the data related to France metropolitan's electricity consumption. The comparative results obtained from this study indicate the superiority of LSTM model compared to other baseline approaches. Authors in Kumar et al. (2018) used LSTM and GRU for forecasting long-time electricity load demands. The electric power consumption in household with a one–minute interval data is used for the training, validation and testing of the stated approaches.

The reader can refer Table 2 for details about the dataset and comparative benchmark models used as well as obtained results with remark in the direction of DL applications in MTSF and LTSF.

4.2. Renewable Energy Generation (REG) forecasting

The output energy prediction of non-conventional/renewable energy generation is very essential for their expansion and development. The improved output energy is always dependent on the input sources. However, the input sources of renewable energy systems (like wind energy systems and photovoltaic systems) have very uncertain and intermittent characteristics. Therefore, intelligent and accurate prediction algorithms are essential for planning and smooth operation of EUI.

4.2.1. Wind power or wind speed prediction

Out of REG systems, the wind energy conversion system (WECS) plays an important role in providing green energy and an alternative option to fossil fuel-based generations. The output of WES is directly associated with wind speed and therefore, an accurate wind speed forecasting (WSF) algorithm will act as an important feature for maximizing power output through a proper operating strategy, capacity planning and load balancing (Santamaría-Bonfil et al., 2016). The output power of a WECS is depended on the speed of wind flow near the wind turbine. The wind speed varies with time, weather and landscape types. The relationship between wind power and wind speed is cubic in nature (nonlinear), which can be analysed from Eq. (1). As a result, a small error in wind speed forecasting can lead to larger errors in wind power prediction (Wani et al., 2020). Therefore, accurate WSF is highly essential for optimal operation and integration of WECS in the power grid.

$$P = \rho A v^3 \tag{1}$$

where, the notation P, ρ, A and v represents wind power (in watt), swept area of wind turbine (m^2), density of air (kg/m^3) and wind speed (m/s) respectively.

The WSF solutions are also classified in accordance with forecasting horizons and time-scales. Fig. 5 shows this classification and portrays the importance of each type (Chen and Folly, 2018). The wind power/speed forecasting methods are also divided into three types: (i) Persistence method (Soman et al., 2010) (It is a benchmark method and highly accurate for STWF Potter and Negnevitsky, 2006), (ii) Physical method (Use of climatological data for example wind speed and direction, pressure, temperature, humidity, terrain structure etc. and very accurate for LTWF) and (iii) Statistical model based such as ML and expert system based (It is simple, low-cost, and provides appropriate predictions for all STWF, MTWF, and LTWF).

Although several intelligent ML and expert systems based WSF methodologies were devolved and presented recently (Liu and Chen, 2019; Zafirakis et al., 2019; Liu et al., 2019), the DL based approaches for WSF problems are shown to be more impactful as per the following referred articles such as Ref. Tascikaraoglu and Uzunoglu (2014), Liu et al. (2019), Wang et al. (2019a), Lytras and Chui (2019) and Marugán et al. (2018). The main aim of this section is to present a detailed review of wind power/speed forecasting considering DL methods.

Authors in Wang et al. (2016b) proposed a hybrid model using wavelet transform, DBN and spine-QR for short-term wind speed forecasting. The wavelet transform (WT) was applied to decompose the natural wind speed time-series data into several frequency series having improved features characteristics. Afterwards, the non-linear features and invariant structures of each frequency were obtained by layer-wise pre-training of deep belief network. Finally, the uncertainties in wind speed were statistically synthesized through the QR technique. Validation of the proposed approach was conducted employing the real-time data extracted from the wind farm located in China (Shangchuan Island wind farm, Guangdong Province) and Australia (Cathedral Rocks wind farm). A comparative result analysis was also conducted between the proposed WT+DBN+QR approach, Auto-Regressive and Moving Average Model (ARMA), the well-tuned Back-propagation Neural Network (BPNN), and the Morlet Wavelet Neural Network (MWNN). The outcome of the study proves the supremacy of the stated approach. Authors in Wang et al. (2017a) suggested a novel STWF model based on WT and CNN. WT was employed for decomposing the raw input data into different frequency band. Then, the nonlinear features from each frequency band were extracted and learned by CNN network to provide an improved prediction result. Similarly,

Liu et al. in 2018 have combined three different approaches empirical-WT, LSTM and Elman neural network (ENN) to forecast wind speed (Liu et al., 2018a). The stated work was accomplished in three steps: (i) empirical-WT was implemented to decompose the original wind speed time-series data into numerous sub-bands with different frequency; (ii) the LSTM network was used to forecast the low-frequency sub-bands, whereas the ENN was used to forecast the high-frequency sub-bands; (c) the forecast outcomes of each sub-band were summed up to acquire the final results for the raw wind speed data. Moreover, other seven types of existing wind speed forecasting (WSF) methods were compared and analysed to show the effectiveness of the suggested approach. Liu et al. proposed a similar types of approach for STWF, where they (Liu et al., 2018b) utilized the wavelet packet decomposition (WPD) techniques to decompose the raw wind speed time-series data into different frequency sub-layer, and then employed CNN with 1D convolution operator to forecast the wind speed high frequency sub-layers and CNNLSTM for forecasting wind speed

Table 2
Summary of some selected DL application in load forecasting.

Proposed model (DL based)	Data sources	Forecasting horizon	Benchmark model	Result		Remark	Ref.
				Benchmark model	Proposed Model		
DRNN–GRU	Residential load data in Austin, Texas, USA (Dataport website) https://dataport.cloud/	1-h resolution 1–2 Month	MLP, ARIMA, SVM, MLR	For instance: MLP: 2.206 (RMSE) 1.672 (MAE) 14.496% (MAPE)	DRNN–GRU: 0.510 (RMSE) 0.345 (MAE) 3.504% (MAPE)	The proposed approach is suitable for forecasting of load demand for short to medium period. The suggested DRNN–GRU model can include time dependencies in the load-data and attain a greater predicting accuracy as compared to the other baseline approaches. The method also proven to be effective for filling the missing values by learning from historical-data, as it can achieve high forecasting accuracy. Nevertheless, the proposed approach has a little restriction like it needs the information of the upcoming climate-data for accurate prediction	Wen et al. (2020)
CNN	Dataset contains 2,075,259 samples measured by a smart meter installed at a house near Sceaux, Paris, France. (Data Retrieved from UCI Machine Learning Repository. https://archive.ics.uci.edu/ml/datasets.html)	15 min	SVM, ANN	SVM: 1.12 (MAE), 1.25 (RMSE), 25.56 (MAPE) ANN: 1.08 (MAE), 1.15 (RMSE), 23.25 (MAPE)	CNN-2D: 0.59 (MAE), 0.79 (RMSE), 12.54 (MAPE)	The performance based on MAE, RMSE, and MAPE criteria presented that the stated method using CBB-2D achieved better result (MAPE of 12.54%) in comparison with 1-D CNN and other traditional ML approaches.	Estebsari and Rajabi (2020)
MLP based DL	Three cities in Jiangsu province, China, (Data source link is not available)	NA	Random Forrest (RF), Gradient boosting tree (GBT)	RF 0.05925 (MAPE %) 0.170209 (MRPE%) 1060.406 (MAE (kWh)) GBT 0.043692 (MAPE %) 0.077298 (MRPE%) 775.1881 (MAE (kWh))	MLP: 0.03550405925 (MAPE %) 0.057273 (MRPE%) 620.0159 (MAE (kWh))	The output results prove that the DL method achieves better than RF and GBT approaches with respect to prediction errors.	Guo et al. (2018)
GA based LSTM	Réseau de Transport d'Électricité (RTE) power consumption data http://www.rtefrance.com/fr/article/bilanselectriques-nationaux	2 Weeks, Between 2–4 Weeks, Between 2–3 Months, Between 3–4 Months	Extra Trees, Ridge Regression, k-Nearest Neighbour, Random Forest, Gradient Boosting, Neural network	For instance of Extra tree: RMSE Extra Trees: 513.8 (Mean), 90.9 (**Std. Deviation**) MAE Extra Trees: 344 (Mean), 55.8 (**Std. Deviation**)	RMSE LSTM: 378 (Mean), 59.8 (Std. Deviation) MAE LSTM: 270.4 (Mean), 45.4 (Std. Deviation)	The proposed approach was found to be suitable for forecasting of load demand in both short and mid-term time horizon. The accuracy of the proposed approach was also compared with many other baseline approaches to prove the efficiency.	Bouktif et al. (2018)

(*continued on next page*)

low frequency sub-layers. Authors (Zhou et al., 2018) proposed a hybrid approach for STWF model based on variational mode decomposition (VMD) and CNN.

Researchers (Chen et al., 2018) proposed a novel technique called EnsemLSTM by using a nonlinear-learning ensemble of DL time series forecast based on LSTMs, support vector regression machine (SVRM) and extremal optimization (EO) algorithm. Initially, with the intention to prevent the disadvantage of low generalization capability and robustness of a single DL approach when handling the diversiform dataset, a group of LSTMs with various hidden layers and neurons are utilized to discover and deed the hidden info of wind speed time-series. Afterwards, predictions of LSTMs were combined into a nonlinear-learning regression top-layer composed of SVRM, and the EO was

Table 2 (*continued*).

Proposed model (DL based)	Data sources	Forecasting horizon	Benchmark model	Result		Remark	Ref.
				Benchmark model	Proposed Model		
GRU	(Data source link is not available)	1 min	RNN, LSTM	RNN:0.623 (MAE) LSTM:0.602 (MAE)	GRU: 0.589 (MAE)	The accuracy of load forecasting with the original set of hyper-parameters is 99.10%. However, the top selected structure of LSTM & GRU has 99.38% accuracy on the test-data-set, which is around 34% decreased testing-error. The result shows that the GRU model was outperform to LSTM and RNN.	Kumar et al. (2018)
SAD	Electricity consumption data of city Los Angeles, New York and Florida (Data source link is not available)	1 h	SVR ANN	SVR: 1.7153 ANN: 3.2878	SAD: 0.9552	Although the output performance (MAPE) of the proposed stack auto-denoising method perform well compared to ANN and SVM, the work need to revalidated with considering a larger sampled dataset with other performance indices.	Tong et al. (2018)
DBN	Australian Energy Market Operator (AEMO) Especially, New South Wales (NSW), Tasmania (TAS), Queensland (QLD), South Australia (SA) and Vic-toria (VIC) (AEMO, Australian Energy Market Operator 2013, 2013. http://www.aemo.com.au/.)	0.5 h 1 h	Persistence SVR ANN RF	For Example: Data: (NSW) Month: January Method: Persistence RSME: 978.24 MAE: 8.55% Method: SVR RSME: 703.43 MAE: 6.23% Method: ANN RSME: 750.53 MAE: 7.2%	Method: DBN RSME: 639.75 MAE: 5.95% Method: EMD+DBN RSME: 541.53 MAE: 4.62%	The DBN based deep learning method is outperform to other baseline method. Moreover, the ensemble deep learning model with empirical mode decomposition has shown more improved result. But, it consume more time as compared to the single structure model.	Qiu et al. (2017)
DRNN pooling-based-DNN (PDRNN)	Electricity Customer Behaviour Trials (CBTs) initiated by Commission for Energy Regulation (CER) in Ireland. (CER, 2011)	0.5 h	ARIMA SVR	ARIMA: 0.5593 (RMSE, 'kWh') 0.2998 (MAE, 'kWh') SVR: 0.5180 (RMSE, 'kWh') 0.2855 (MAE, 'kWh')	DRNN: 0.4815 (RMSE, 'kWh') 0.2698 (MAE, 'kWh') PDRNN: 0.4505 (RMSE, 'kWh') 0.2510 (MAE, 'kWh')	The proposed PDRNN method was outperform to several other baseline method. The proposed PDRNN have shown an improved result (RMSE & MAE) with respect to ARIMA and DRNN with the following value respectively: (19.45% & 16.28%) and (6.45% & 6.96%).	Shi et al. (2017)
DNN, DNN–RNN Parallel-CNN–RNN	Load values of a city in North China. (Data source link is not available)	Hour-ahead Day-ahead Week-ahead	LR, SVR	LR: 2.761 (% MAPE), SVR: 1.650 (% MAPE),	DNN: 2.761 (% MAPE), DNN–RNN: 1.650 (% MAPE), Parallel-CNN–RNN: 1.349 (% MAPE),	The hybrid CNN–RNN approach have shown better result compared to other baseline ML model and also some other DL method such as DNN and RNN.	He (2017)
DNN	90-days hourly data of Iberian Energy Market Operator (MIBEL). http://imagenet.org/challenges/LSVRC/2016/index	1 day	NA	NA	DNN with ReLU activation function: 1.3 (% MAPE), DNN with ELU activation function: 1.52 (% MAPE),	In this work, authors have used multiple activation function and its combination to train the DNN. Although the approach perform nicely to predict the one-day ahead load data of MIBEL data, the comparative analysis with other baseline approaches are missing in this paper.	Hossen et al. (2017)

(continued on next page)

presented to optimize the parameters of the top-layer. Finally, the final ensemble WSF was specified by the fine-tuning top-layer.

Zhu et al. investigated the issues of forecasting wind speed for multiple sites simultaneously and provided a solution/prediction model

Table 2 (*continued*).

Proposed model (DL based)	Data sources	Forecasting horizon	Benchmark model	Result		Remark	Ref.
				Benchmark model	Proposed Model		
DBN	urbanized area in Texas, United States (Data source link is not available)	Day-ahead, Week-ahead	NN, SVR, ELM,	For example (In Summer and day-ahead): NN: 6.36 (% MAPE), 1729.13 (RMSE) SVR: 6.52 (% MAPE), 1734.81 (RMSE) ELM: 6.55 (% MAPE), 1735.96 (RMSE)	For example (In Summer and day-ahead): DBN: 5.22 (% MAPE), 1517 (RMSE)	Although the DBN based DL approach was performed well in term of RMSE and MAPE to forecast the load demand, the author have integrated the Copula model with DBN to decrease the error further.	He et al. (2017)
LSTM	Data collected from Smart Grid Smart City project initiated by the Australian Government http://www.industry.gov.au/ENERGY/PROGRAMMES/SMARTGRIDSMRTCITY/Pages/default.aspx	Day-ahead, Week-ahead	BPNN, KNN, ELM	For example (Average individual forecast day-ahead): BPNN: MAPE of 49.49%. KNN: MAPE of 81.13%. ELM: MAPE of 123.45%	LSTM: MAPE of 44.06%	In this work, author have carried out the load forecasting task with different time step such as 2,6,12 etc. Here it can be noticed that although individual load forecasting is far good than aggregating all individual forecasts. Moreover, the it can be noticed that the LSTM shows better result compared to the conventional strategy of directly forecasting the aggregated load.	Kong et al. (2019)

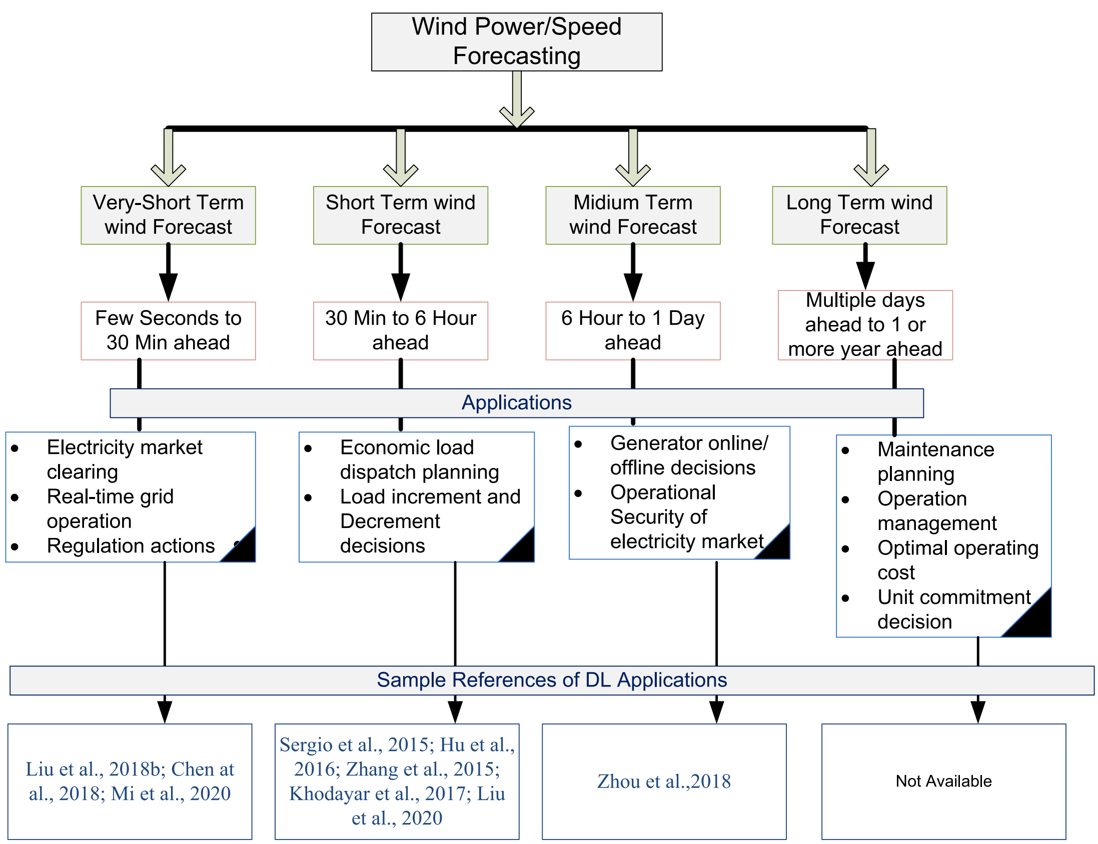


Fig. 5. Taxonomy of Wind power/ speed forecasting methods based on time horizons.

for the same (Zhu et al., 2018). The author's suggested WSF model was based on spatiotemporal correlation i.e. predictive-DCNN. The experimental outcomes via handling the real-time dataset confirmed that predictive-DCNN was able to capture the spatiotemporal correlation efficiently, and it outperforms the traditional ML tools, comprising MLP, SVR, DT, etc.

A data-driven multi-model WSF procedure was suggested in Feng et al. (2017). The stated approach was based on a two-layer ensemble ML method. The initial layer was comprised of manifold ML models that provide distinct predictions. A deep FS framework was built to decide the best suitable inputs to the initial layer ML models. Subsequently, a blending method was operated in the succeeding layer to

Table 3
Summary of some selected DL application in wind power forecasting.

Proposed model (DL based)	Data sources	Forecasting horizon	Benchmark model	Result		Remark	Ref.
				Benchmark model	Proposed model		
DBN	Northeastern region of Brazil (Source data link is not available)	1 h	MLP	**For wind farm-1 (BJD)** 0.6190 (MSE), 0.6068 (MAE) **For wind farm-5 (TRI)** 1.815 (MSE), 1.011 (MAE),	**For wind farm-1 (BJD)** 0.6731 (MSE), 0.632, (MAE), **For wind farm-5 (TRI)** 1.52 (MSE), 0.935 (MAE),	Although in some dataset the proposed model shown to be outperforming to benchmark model,there was no strong evidence about the pre-eminent method for each datasets.	Sergio and Ludermir (2015)
SVR, ELM	Four wind Farms are located in Ningxia, Jilin, Inner Mongolia and Gansu, respectively. (Source data link is not available)	10 min, 30 min, 1 h, 2 h	SHL-DNN	**For example, 10-min Forecasting horizon** SVR: 0.8545 (MAE) 20.23 (MAPE) ELM: .6578 (MAE) 12.33 (MAPE)	**For example, 10-min Forecasting Horizon** 0.5348 (MAE) 10.35 (MAPE)	The suggested wind forecasting model shown as effective one for different predication horizon. However, it can be noticed from the result that a significant reduction in the performance of prediction rate as a result of increasing forecasting horizon.	Hu et al. (2016)
CNNLSTM	Xinjiang Uygur Autonomous Region, northwest of China (Source data link is not available)	10–30 min	ARIMA, SVM, BP, RBF, ENN, ELM	**For example, (ARIMA):** 10.07 (MAPE), 1.07 (MAE), 1.34 (RMSE) **(ELM):** 10.89 (MAPE), 1.21 (MAE), 1.70 (RMSE)	5.07 (MAPE), 0.58 (MAE), 0.37 (RMSE)	The method can perform better compared to other benchmark models under extremely disturbed condition (for example, the wind speed experiences impulsive variation).	Liu et al. (2018b)
Predictive-DBM (PDBM)	Southern China(Source data link is not available)	10 min–120 min	AR, ANFIS, SVR	**For example, 10-min Forecasting horizon** AR: 0.3451 (MAE) 7.57 (MAPE) ANFIS: 0.34 (MAE) 10.84 (MAPE) SVR: 0.6340 (MAE) 14.55 (MAPE)	0.2951 (MAE), 7.05 (RMSE)	One open problem for the PDBM model is that it necessitates longer learning time, particularly when it has more hidden feature layers and more hidden units. On the other hand, the larger scale PDBM possibly will offer improved interpretation.	Zhang et al. (2015)
VMD–CNN	Inner Mongolia, China and Sotavento Galicia, Spain (Lydia et al., 2016)	16 h	SVR, ELM	SVR: 0.8545 (MAE) 20.23 (MAPE) ELM: .6578 (MAE) 12.33 (MAPE)	SVR: 0.8545 (MAE) 20.23 (MAPE) ELM: .6578 (MAE) 12.33 (MAPE)	The suggested VMD–CNN based model has shown improved percentage of 72.22 and 79.33 compared to SVR and ELM respectively, regarding RMSE. Similarly, 74.12 and 80.88 improved percentage shown by stated method compared to SVR and ELM respectfully, regarding MAE.	Zhou et al. (2018)

(*continued on next page*)

generate a group of the predictions given by primary layer models and produce both deterministic and probabilistic predictions. This two-layer model seeks to employ the statistically dissimilar characteristics of each ML technique.

Table 3 (*continued*).

Proposed model (DL based)	Data sources	Forecasting horizon	Benchmark model	Result		Remark	Ref.
				Benchmark model	Proposed model		
DBN	Shangchuan Island wind farm and the Cathedral Rocks wind farm (Wan et al., 2013)	1 h	ARMA, BPNN, MWNN	In comparison to Compared to ARMA, BPNN, and MWNN, the MAE index has been improved by 44.18%, 48.99%, and 48.47%, respectively.		Some cream points related to this model are (i) the performance of the model is insensitive to number of simulation (ii) the model have a invariant structure for high level feature extraction	Wang et al. (2016b)
RSDAE	Idalia, Colorado (Source data link is not available)	10 min, 30 min, 1–3 h	PR, FFNN, TDNN, NARNN	**For example, 10-min Forecasting horizon** **PR:** 0.317 (MAE) 0.631 (RMSE) **FFNN:** 0.308 (MAE) 0.627 (RMSE) **TDNN:** 0.298 (MAE) 0.619 (RMSE) **NARNN:** 0.281 (MAE)0.603 (RMSE)	**For example, 10-min Forecasting horizon** RSDAE: 0.317 (MAE) 0.631 (RMSE)	The suggested model shown to be more effective compared to other benchmark model (for example, the average improved performance exhibited by the stated approach was 13.6%–4.5% (RMSE), 24.2%–11.2% (MAE)). However, it can be noticed from the result that a significant reduction in the performance of prediction rate as a result of increasing forecasting horizon.	Khodayar et al. (2017)
LSTM	Inner Mongolia, China (Source data link is not available)	10 min	ARIMA, SVR, KNN, GBRT	**ARIMA:** 1.37 (MAE), 20.73 **(MAPE),** 1.83 (RMSE) **SVR:** 1.18 (MAE), 17.75 (MAPE), 1.57 (RMSE) **KNN:** 1.22 (MAE), 17.92 (MAPE), 1.62 (RMSE)	**LSTM:** 1.14 (MAE), 17.10 (MAPE), 1.53 (RMSE)	The suggested model shown to be more effective compared to other benchmark model (for example, the average improved performance exhibited by the stated approach were 8.74% (MAE), 1.36% (MAPE)).	Chen et al. (2018)
LSTM	Dataset 1: real-life data from Sotavento that is located in the south-west of Europe, in Galicia, Spain, Dataset 2: Kerman that is located in the Middle East, in the southeast of Iran (http://www.sotaventogalicia.com/)	1 h	MLP	**For instance:** **Dataset 1:** **MLP:** 2.6867 (MAE) 3.4945 (RSME) 38.4705 (MAPE) **Dataset 2:** **MLP:** 2.1418 (MAE) 2.6213 (RSME) 70.7681 (MAPE)	**Dataset 1:** **LSTM:** 2.0018 (MAE) 2.6535 (RSME) 28.0287 (MAPE) **Dataset 2:** **LSTM** 1.8017 (MAE) 2.1735 (RSME) 57.4884 (MAPE)	The authors have shown a novel method to forecast the wind power using LSTM. The result presented in this table is only a comparison of DL based and ML based method extracted from the paper. However, the actual proposed method in this paper is based on the integration of WT, PSO based feature selection and optimization, and LSTM. The hybrid approach has shown improved result compared to MLP and LSTM like MAE: 0.1217, RMSE: 0.1536, MAPE:4.0857.	Memarzadeh and Keynia (2020)

(*continued on next page*)

Table 3 (*continued*).

Proposed model (DL based)	Data sources	Forecasting horizon	Benchmark model	Result		Remark	Ref.
				Benchmark model	Proposed model		
STNN	Wind form located in USA (NREL website) https://www.nrel.gov/grid/wind-toolkit.html	1 h, 2 h, 3 h	ANN, LR, CNN, LSTM	For instance: (1 h Forecast horizon) RMSE: 1.865 (ANN), 1.866 (LR), 1.826 (ANN), 1.832 (ANN)	For instance: (1 h Forecast horizon) RMSE: 1.714 (STNN),	The suggested spatial–temporal neural network (STNN) pools the convolutional-GRU model and 3D-CNN and uses an encoding-forecasting structure to generate the spatiotemporal predictions. Variational Bayesian inference is engaged to acquire the approached posterior parameter distribution of the model and determine the probability of the prediction. The investigational outcomes prove that the suggested approach ominously outperforms compared to other baseline approaches.	Liu et al. (2020)
LSTM LSTM–GRU	Four wind farm sites in the United State (collected from Global monitoring laboratory website) http://www.esrl.noaa.gov/gmd/grad/surfrad/	1 h, 2 h, 3 h	ARIMA, BP, Elman, ELM,	For instance: (1 h Forecast horizon) RMSE: 0.5496 (ARIMA), 0.4141 (Elman), 0.6405 (BP), 0.4471 (ELM)	0.4057 (LSTM) 0.4727 (LSTM–GRU) 0.3806 (GRU)	The author have presented a detail analysis of DL based wind power forecasting and compared their work with several ML approach based on three time steps horizon and several error indices such as RMSE, MAE and MPAE. This table shows an instance of 1 h time horizon prediction result based on RMSE obtained. The result shows that the suggested approach GRU approaches outperform to other method with less error and computational time.	Peng et al. (2020)
ESN, DeepESN	collected from the China Statistical Yearbook 2018 http://data.stats.gov.cn/easyquery.htm?cn=E0101	1 h.	HFGSE, EMD-LSSVR, GM-ARIMA, persistence, BPNN	For instance: (1 h Forecast horizon) RMSE: 0.2341 (HFGSE), 0.3907 (EMD-SSVR), 0.3152 (GM-ARIMA) 0.4264 (persistence), 0.1819 (BPNN)	For instance: (1 h Forecast horizon) RMSE: 0.1401 (ESN) 0.0701 (DeepESN)	In the comparative result analysis the DeepESN shows the best forecasting performance compared with all the other baseline approaches in regards of RMSE, MAPE and computational efficiency.	Hu et al. (2020)

(*continued on next page*)

Authors in Wang et al. (2017a) have proposed a new DL based ensemble approach for probabilistic wind power forecasting (WPF). The suggested approach was based on WT and CNN. The WT was implemented to decompose the original wind power time series data into

Table 3 (*continued*).

Proposed model (DL based)	Data sources	Forecasting horizon	Benchmark model	Result		Remark	Ref.
				Benchmark model	Proposed model		
SSA-MADANET	Four wind form (WF) from Xinjiang, China (Source data link is not available)	10 min 20 min 30 min	ARIMA ADANET MADANET	**For instance:** **WF-1: (10 min)** ARIMA: 7.41% (MAPE), 0.75 (MAE), 1.02 (RMSE) ADANET: 7.19% (MAPE), 0.73 (MAE), 1.03 (RMSE)	**For instance:** **WF-1: (10 min)** MADANET: 7.05% (MAPE), 0.71 (MAE), 1 (RMSE) SSA-MADANET: 3.51% (MAPE), 0.35 (MAE), 0.46 (RMSE)	The ADANET is a novel and capable data-dependent learning and DL algorithm to forecast the wind power. It is appropriately used for capturing the vibrant features in wind speed. This work has modified the ADANET to a new MADANET by combining it with the MS layer and LSTM network. The stated approach can obtain strong and high-precision WPF.	Mi and Zhao (2020)
WT+DBN	The four wind time-series speed data (s1–s4) collected from Sichuan Province, China (Source data link is not available)	NA	ENN LGBM RF	**For instance:** **WF-s1: (1 Step)** ENN: 46.13% (MAPE), 1.4211 (MAE, m/s), 1.9629 (RMSE) RF: 53.78% (MAPE), 1.4807 (MAE), 2.0174 (RMSE)	**For instance:** **WF-s1: (1 Step)** WT+DBN: 0.9117 (MAE, m/s) 22.79% (MAPE) 1.0739 (RMSE)	In this work several ML based and DBN based DL method integrated with or without WT based feature extraction technique are tested to predict the wind power. The WT+DBN method shows to be better approach compared to others. Moreover, it can be analysed that less than 10 min required training the network and can be useful for real-time prediction.	Jiajun et al. (2020)
CEEMDAN-error-VMD-LSTM	NREL National Wind Energy Technology Center (NWTC) (Jager and Andreas, 1996) https://www.nrel.gov	15 min, 1 h	MLP, SVR, LSTM, EEMD-LSTM CEEMDAN-LSTM	**For instance forecasting of 15 min horizon** **RMSE:** 0.88588 (MLP) 0.8784 (SVR) 0.8729 (LSTM) 0.5212 (EEMD-LSTM) 0.3427 (CEEMDAN-LSTM)	**For instance forecasting of 15 min horizon** CEEMDAN-error-VMD-LSTM 0.1110 (RMSE) 3.21% (MAPE) 0.0831 (MAE, m/s)	Here, a hybrid approach of WPF based on error correction, double decomposition (DD) and DL is suggested. The DD policy is proved to have better result compared to single decomposition. Considering series decomposition in error correction models, the proposed approach provides better prediction accuracy than the models without error series decomposition. In particular, the CEEMDAN-error-VMD-LSTM model reduced the MAPE by 62.20%, 57.80%, 66.66% and 69.88% than the CEEMDAN-error-LSTM model in four experimental cases.	Ma et al. (2020)

different frequencies sub-layers. Afterwards, each sub-layer was feed as input to CNN to learn deep features automatically and effectively. The forecasting accuracy obtained by the suggested approach was later compared with the benchmark persistence technique and shallow NN models, such as BP and SVM.

Torres et al. (2018) presented a predictive model based on deep-MLNN and meteorological data for predicting the wind farm generation for the next 24 h. This model was trained and validated with a dataset collected from a wind farm located at the island of Tenerife. Yu et al. (2019) presented an improved LSTM-enhanced forget-gate network model (LSTM-EFG) for short-term WPF. Based on the correlation, the feature dataset collected from groups of turbines separated from each other with a certain distance was filtered to further optimize the forecasting effect on wind power by clustering. The prediction outcome indicates that the technique with Spectral Clustering has a higher accuracy with an increase of 18.3% than those of the other predicting models, and at the same time the convergence process has been sped up.

Wu et al. (2016) proposed a novel method based on DNNs for the deterministic short-term WPF. DNNs model including LSTM–RNN has achieved improved results compared to conventional predicting models. In addition to this, a probabilistic WPF based on conditional error analysis was also executed in the stated work. The performance of the stated technique was verified on a data-set comprising data from numerous wind farms in north-east China. Wang et al. (2018b) proposed a DBN model based STWF. The numerical weather prediction (NWP) data was carefully chosen as the input of the suggested model. The NWP dataset and wind data in the wind generating station have shown to be similar features. Thus, the authors have used *k*-means clustering algorithm to deal with the uncertainty in NWP dataset. Afterwards, the output of clustering analysis which provide a large number of NWP samples, were feed as input to DBN model for getting an improved prediction rate. The DBN model was authenticated by the Sotavento wind generating station in Spain.

Qureshi et al. (2017) presented a novel STWF model with the help of DNN based ensemble method with transfer learning concept. Here, the deep-auto-encoder and DBN were used as base regressor and meta-regressor respectfully. The DBN was formed by stacking RBMs. The stated STWF model was utilized the learning abilities of a group of regressors thus, it was more robust in comparison to other similar approach. Sergio et al. studied the applicability of DL architecture such as DBN, SAE and DSAE in forecasting the hour-ahead average speed of winds in the Northeastern region of Brazil (Sergio and Ludermir, 2015). Hu et al. (2016) presented a SDAE and created a shared-hidden-layer architecture based on transfer learning scheme. The input and hidden layers were distributed for 4 wind generating stations, and the output layers were parted to obtain the distinct prediction outcome of each WES. The stated transfer learning technique was able to use mutual features in better ways which were assembled from different WES. With an intention to explain the uncertainty of wind flow, Khodayar et al. (2017) combined rough-NNs with SAE and SDAE by using rough neurons in the hidden and output layer. The suggested models were called as rough-SAE (RSAE) and rough-DAE (RSDAE). Implementation outcomes showed improved generalization capability and predicting performance compared to other standard prototypes, such as SAE and SDAE.

Huang et al. (2018) proposed an improved innovative STWF model based on ensemble-EMD method and hybrid forecasting technique comprising Gaussian process regression (GPR) and the LSTM. They used EEMD to decompose the original raw wind time-series data into several IMFs and then, the LSTM and GPR approach were applied to forecast the IMFs respectively. Afterwards, analysing the IMFs' prediction results from these two predictors, the variance–covariance method was used to calculate the weight of these two predictors and provide a combine result.

Zhang et al. (2015) presented a complex DL architecture model for STWF and LTWF that is, the predictive-DBM (PDBM) and corresponding learning system. The suggested approach predicts wind speed by analysing the advanced level features extracted from lesser level features of the wind speed dataset. These inevitably cultured features were very edifying and suitable for the forecast. The assessment of the offered PDBM model was portrayed by both hourly and daily basis forecast study using real-time wind speed dataset. The forecast accuracy of the stated PDBM model outperforms prevailing approaches by in excess of 10%. Hu and Chen (2018) presented an innovative non-linear hybrid approach pointing at advancing forecast result of wind speed named LSTMDE–HELM based on LSTM, Hysteretic Extreme Learning Machine (HELM), Differential Evolution algorithm (DE), and nonlinear collective mechanism. The suggested non-linear LSTMDE–HELM system has been exercised on the dataset collected from a wind generating station situated at Mongolia, China. The forecasting horizon chosen for these experiments were ten-minutes ahead and hour-ahead. Table 3 describes the findings of some randomly selected articles based DL models in wind energy predicting application. It can be concluded that although the application of DL in WEF is not fully-explored yet, the predicting performances are mostly untouchable compared to superficial predictors, irrelevant to areas and forecasting horizons. However, the DL based architecture forgoes computational time to attain greater precision.

4.2.2. PV power forecasting

PV generating system (PVGS) has gained huge attention by the global EUI and therefore, its rate of integration into the main-grid is increasing day-by-day. According to the global energy status report (REN21) accessible from REN21 (2018), the global PV systems installed power capacity has increased immensely from 8 GW in 2007 to 402 GW in 2017. The output energy of PVGS is highly dependent on solar irradiance, temperature, and different weather constraints. Thus, to affirm reliable operation and economic integration of PVGS in SG, appropriate forecasting of photovoltaic generation is an essential subject for research (Ahmed et al., 2020; Mellit et al., 2020). Similar to load forecasting and wind speed forecasting problem, and its solution as presented in Sections 4.1 and 4.2.1 respectively, the PV forecasting (PVF) methods can also be distributed on different time horizons. Fig. 6 shows this classification and portrays the importance of each type (Akhter et al., 2019). The PVF methods are generally classified as persistence methods, statistical methods, ML-based methods, and hybrid methods. Akhter et al. presented a nice and comprehensive review of each PVF method. The authors have focused mainly on the ML methods for PVF and carefully reported its importance in this field. Wang et al. provided a comprehensive review on AI applications to advanced solar power forecasting (Wang et al., 2020). As we have already stated that DL methods belong to the advanced version of AI that is AI-2.0 and therefore, the research and application of DL in PVF have been increased in recent years.

In this section, we have tried to provide a detailed overview of the status and scope of DL in PVF which may definitely help the EUI researches those who are working in this specific area.

Nasser and Mahmoud in Abdel-Nasser and Mahmoud (2019) proposed an LSTM and RNN based method for forecasting the output power of PVGS. The stated approach was tested and evaluated on two hourly PV datasets collected from two different cities of Egypt such as Aswan and Cairo. A comparative performance analysis was also carried out with multiple linear regressions (MLR), bagged regression trees (BRT) and NN methods. The result shows that the proposed approach was outperformed with very little RMSE. Gensler et al. in Gensler et al. (2016) used the different DL framework for short-time solar power forecasting. In this framework, they have implemented a physical forecasting model, MLP, LSTM, DBN and auto-encoders for predicting the solar power of 21 PV facilities located in Germany. A comparative analysis was also carried out based on the output accuracy.

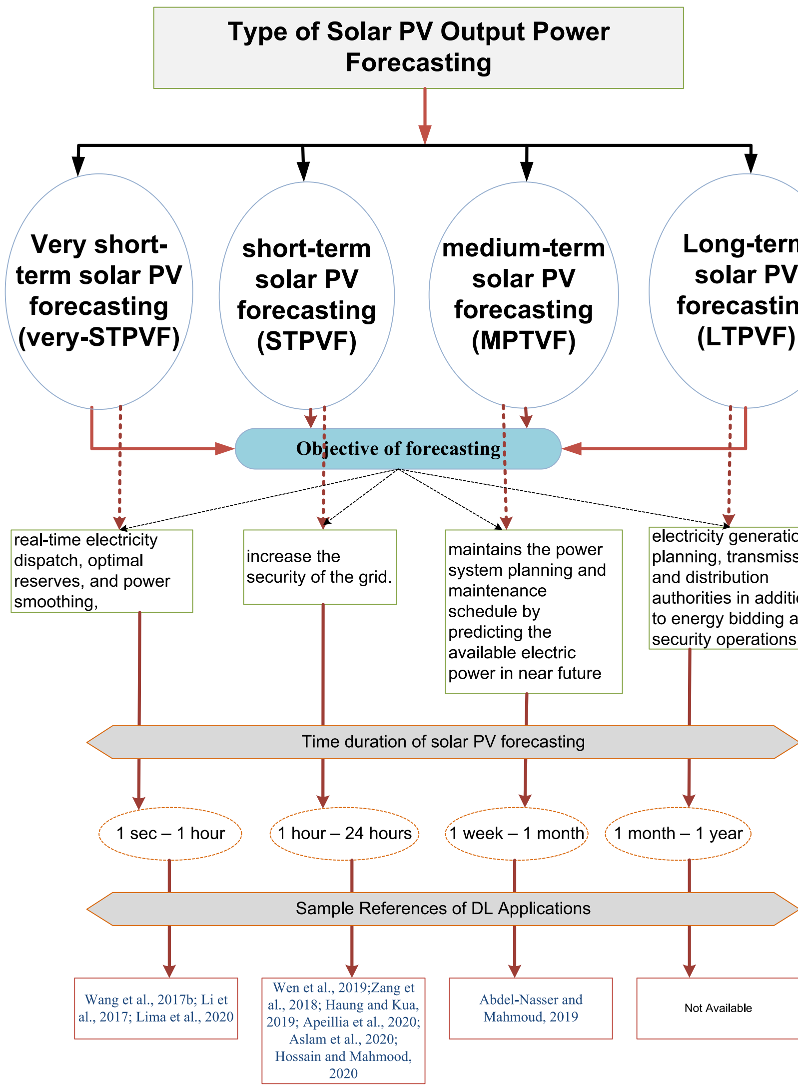


Fig. 6. Classification of PV output power forecasting based on time horizon (Akhter et al., 2019).

In Wen et al. (2019), Wen et al. have proposed a DRNN–LSTM based SPV power load forecasting model and proved that it performs well compared to traditional MLP and SVM based tools.

By using meteoroidal data (for example, temperature and solar irradiance) and historical PVGS output data as input to train and test DNNs, Haung and Kua in Huang and Kuo (2019) proposed a high precision DNN model named "PVPNet" to forecast 24 h PV system output power. Li et al. in Li et al. (2017) used DBN for predicting short time output power PV generation. The input variables to the DBN model were solar irradiation, atmospheric pressure, relative humidity and past 4 days output solar power data.

A few papers have shown the advantage of using decomposition methods on historical PV output time series data, which are later used as input to train the deep neural network model. In this framework, Wang et al. proposed a hybrid model based on wavelet transform and DCNN on PV output time series data collected from PV-farms in Belgium (Wang et al., 2017b). The computational results indicate that the mean MAPE, RMSE, and MAE of the offered deterministic model outperform the three other compared benchmarks models in terms of seasons, forecasting horizons and PV power locations. Similarly, the author in Zang et al. (2018) presented VMD and CNN for forecasting PV output power on a daily and hourly basis. The sky image data and surface irradiance measurement were used by CNN and LSTM model to SPVF in Zhen et al. (2020). The performance of the stated approach provides higher accuracy and can preserve robustness under different weather conditions.

Solar irradiance is one of the most significant meteorological data, which mainly responsible for the change in PV output power. Therefore, accurate solar irradiance prediction will definitely help to predict accurate solar output power. Alzahrani et al. in Alzahrani et al. (2017) proposed a deep-RNN model for predicting solar irradiance levels on an hourly basis in a day. Table 4 describes the findings of some randomly selected articles based on DL models in solar energy predicting application.

4.3. Power system state forecasting

Power system state forecasting (PSSF) performs an essential character in electrical utility system monitoring, by presenting system awareness even in advance of the time horizon, improving system observability, and offering effective recognition of the grid topology and link parametric changes. The rapid voltage fluctuation in the electrical utility grids is becoming a major challenge as a result of the high penetration of distributed energy resources; electric vehicles, energy systems, and cyber assist system into the main grid. Therefore,

Table 4
Summary of some selected DL application in solar power forecasting.

Proposed model (DL based)	Data sources and input variable	Forecasting horizon	Benchmark model	Result		Remark	Ref.
				Benchmark model	Proposed model		
LSTM–RNN	Aswan (Dataset1) and Cairo (Dataset2) cities, Egypt. (Source data link is not available)	Hour-ahead, day-ahead, week-ahead	MNL, BRT, NN	**MNL:** 384.89 (RMSE) **BRT:** 494.46 (RMSE) **NN:** 377.072 (RMSE)	**LSTM–RNN:** 82.15 (RMSE)	***Input variable: Historical Solar data*** The suggested approach has very small prediction error related to benchmark model. However, the methods have considered only one error indices RMSE" for comparisons purpose.	Abdel-Nasser and Mahmoud (2019)
DRNN–LSTM	Sails in the Desert, Yulara, Australia (Source data link is not available)	Hour-ahead,	MLP SVM	**MLP:** 12.181 (RMSE), 7.526 (MAE), 34.71% (MAPE) **SVM:** 14.92 (RMSE), 11.549 (MAE), 38.96% (MAPE)	**DRNN–LSTM:** 7.53 (RMSE), 4.369 (MAE), 15.87% (MAPE)	***Input variable: Historical Solar data*** The suggested approaches have shown to be highly accurate compared to the other contemporary models with less error index.	Wen et al. (2019)
DBN	The Desert Knowledge Australia Solar Center (DKASC). http://dkasolarcentre.com.au/	Hour-ahead,	BPNN	**BPNN:** 16.74 (MAPE), 1.0 (MAE), 0.166 (TIC)	**DBN:** 8.92 (MAPE), 0.92 (MAE), 0.039 (TIC)	***Input variable: Historical PV power data, solar radiation intensity, temperature, relative humidity, and wind speed.*** The prediction performance highly dependent on the sunshine hours for example, in summer session, there will be no significant difference between the performances regarding different months but in cloudy and rainy weather have big influence on prediction accuracy.	Li et al. (2017)
WT+QR+DCNN	PV farms in Belgium(north-western Flanders PV farm and north-eastern Limburg PV farm) http://www.elia.be/en/grid-data/power-generation/Solar-powergeneration-data/Graph	15 min, 30 min, 30 min, 90 min, 120, min	BPNN SVM WT+SVM	**For example, 45 min time horizon, BPNN:** 0.0933 (MAPE), 8.1205 (RMSE), 4.6912 (MAE) **SVM:** 0.0748 (MAPE), 7.5710 (RMSE), 4.124 (MAE)	**For example, 45 min time horizon, DCNN:**0.0385 (MAPE), 3.8772 (RMSE), 2.0340 (MAE)	***Input variable: Historical Solar data*** The numerical outcomes indicate that the mean MAPE, RMSE and MAE of the suggested model beat the compared benchmarks models in terms of seasons, forecasting time horizon and solar farm locations.	Wang et al. (2017b)
VMD–CNN	Electric power company in Jiangsu Province, China. (Source data link is not available)	1 h, 6 h, 12 h	Shallow BPNN, SVR, GPR	**For example, 1 hour-ahead prediction, BPNN:0.475** (MASE), 5.77 (RMSE), 4.18 (MAE) **SVR:** 0.450 (MASE), 5.52 (RMSE), 3.96 (MAE) **GPR:0.4029** (MASE), 5.61 (RMSE), 3.54 (MAE)	**For example, 1 hour-ahead prediction, BPNN:** 1.5418 (MAPE), 2.0533 (RMSE), 0.1752 (MAE)	***Input variable: Historical Solar data*** Here, the paper presents a hybrid 2D-VMD–CNN. It attains greater prediction accuracy compared to other benchmark models including 1D VMD-based forecasting approach.	Zang et al. (2018)

(*continued on next page*)

monitoring and effective tracking of system states grow into more and more critical, not only for utility protection (Wang et al., 2019b) but also for energy management (Zhang et al., 2016). The available methods are based on linear estimators or FNN. However, these conventional methods are not able to restrain long-term nonlinear dependencies in the voltage–time series and lead to weak performance. To overcome these disadvantages of conventional state estimators, the current research attention diverted towards the envelopments of DL/RL

Table 4 (*continued*).

Proposed model (DL based)	Data sources and input variable	Forecasting horizon	Benchmark model	Result		Remark	Ref.
				Benchmark model	Proposed model		
Deep-CNN	In a particular place of Taiwan (Source data link is not available)	24 h	DT, RF, SVM, MLP	**DT: 206.61 (RMSE), 140.30 (MAE) RF:167.52 (RMSE), 116.00 (MAE) SVM:185.22 (RMSE), 147.30 (MAE) MLP:224.99 (RMSE), 196.68 (MAE)**	163.15 **(RMSE), 109.48 (MAE)**	***Input variable: Historical PV power data, solar radiation intensity, temperature.*** The suggested model shown to be more effective compared to other benchmark model. Moreover, it can be observed from the result that the proposed PVPNet algorithm can reduce the monitoring expenses, the primary cost of hardware mechanisms, and the Long-term maintenance costs of the future PV farms.	Huang and Kuo (2019)
CNN LSTM	From a500 kWp solar plant located in the south of Taiwan (Zhong et al., 2018)	Day-ahead	SVM	**SVM: 30.66% (MAPE), 4.59% (MRE), Computation time (min): 2.74**	**CNN: 29.72% (MAPE), 2.94% (MRE), Computation time (min): 15.6 LSTM: 35.85% (MAPE), 5.99% (MRE), Computation time (min): 6.14**	***Input variable: Historical PV power data and 5 weather variable.*** The depicted values in this table on the result section are the performance evaluation in rainy weather. It can be observed that the CNN approach shows better result whereas the computational burden is more.	Aprillia et al. (2020)
LSTM, and GRU	Hourly and daily solar radiation data collected from the https://web.kma.go.kr/eng/index.jsp	1 h, 1 day	SVR, RNN, FFNN,	**RMSE: (hourly)** 0.3990 (SVR) 0.3928 (FFNN) **RMSE: (daily)** 5.3618 (SVR) 5.4492 (FFNN)	**RMSE: (hourly)** 0.3920 (LSTM) 0.3909 (GRU) **RMSE: (daily)** 5.3696 (SVR) 5.3315 (FFNN)	***Input variable: Historical Solar data*** The depicted values in this table on the result section are belongs to the power plant located in Seoul. The error index (RMSE) have not shown much difference.	Aslam et al. (2020)
LSTM	Datset1: Brazilian data Datset2: Spanish data (Photovoltaic Geographical Information System (PVGIS) of the European Commis-sion/Institute for Energy and Transport (IET))	hourly	MLP RBF SVR	**For instance:** (Spanish data with Winter season) **MAPE: 8.16% (MLP) 7.22% (RBF) 8.12% (SVR)**	**For instance:** (Spanish data with Winter season) **MAPE: 7.19% (LSTM)**	***Input variable: Solar irradiance and air temperature*** Result shows that the use of deep learning based LSTM model for SPVF has a significant improvement compared to other ML based approaches	Lima et al. (2020)
LSTM	National renewable energy laboratory (NREL) https://www.nrel.gov	0.5 h 1 h 6 h 12 h 24 h	NA	**NA**	**For instance:** (Winter season) **LSTM: 0.61 (MAE) 1.39 (RMSE)**	***Input variable: Historical solar data with weather information.*** Although the performance of the LSTM approach shows a good result on solar forecasting, the stated approach need to compare with other traditional ML approaches.	Hossain and Mahmood (2020)

(*continued on next page*)

based PSSF methods which are able to capture long-term dependencies (Zhang et al., 2019a) and ease to realize. Zhang et al. adopted DRNN and prox-linear nets (RPLN) for PSSF using past measurements (Zhang et al., 2019a). The simulation results indicate that the suggested RNN

Table 4 (*continued*).

Proposed model (DL based)	Data sources and input variable	Forecasting horizon	Benchmark model	Result		Remark	Ref.
				Benchmark model	Proposed model		
LSTM, GRU, RNN	Collected from DKASC, Alice Springs, Australia http://dkasolarcentre.com.au/	1 h	MLP	**MAPE (%):** 10.1575 (MLP) **RMSE (kW):** 11.0861 (MLP)	**MAPE (%):** 7.5978 (LSTM) 8.5169 (GRU) 8.7263 (RNN) 10.1575 (MLP) **RMSE (kW):** 1.0382 (LSTM) 1.0351 (GRU) 1.0581 (RNN) 1.0861 (MLP)	***Input variable: Historical solar data*** The authors have shown the effective ness of LSTM on solar forecasting with respect to MAPE and RMSE. However, they have analysed that the performance may show better result with the integration signal decomposition technique like WPD to LSTM.	Li et al. (2020)
Residual network (ResNet) and dense convolutional network (DenseNet)	Collected from DKASC, Alice Springs, Australia http://dkasolarcentre.com.au/	Day-ahead	SVR, RF, MLP	**SVR:** 0.175 (MSE) 0.272 (MAE) 0.456 (MASE) **RF:** 0.235 (MSE) 0.299 (MAE) 0.501 (MASE) **MLP:** 0.138 (MSE) 0.215 (MAE) 0.360 (MASE)	**ResNet:** 0.128 (MSE) 0.180 (MAE) 0.301 (MASE) **DenseNet:** 0.081 (MSE) 0.152 (MAE) 0.255 (MASE)	***Input variable: Historical PV power series, historical meteorological elements and weather type predictions are utilized as inputs.*** Here, the authors have used two version of CNN such as **ResNet** and **DenseNet**. The forecasting accuracy and reliability of the proposed hybrid deep networks have been through various error index and shown to be great feasibility for practical applications.	Zang et al. (2020)

and EPLN based predictors perform better compared to FNN and support vector auto-regression method.

5. Automatic Power Quality (PQ) monitoring

Deterioration risks in electric PQ have been increased due to the excessive use of power semiconductor devices in a grid utility. PQ disturbances (PQDs) are the results of deviation/disruption in the voltage, current and frequency signals from the benchmark rating (Mahela et al., 2015), which leads to malfunctions of the control system as well as decreases the life span of electrical equipment's. Therefore, PQ management is an important topic for EUI which signifies 'maintaining quality of power in a prescribed limit'. In the corporate world, a trendy axiom says that "you cannot manage what you do not measure". Thus, automatic detection, measurement, and recognition of PQDs are the basis to deal with the PQ problem. In general, feature extraction, selection, and classifications are three important stages of automatic PQDs recognition (Mishra, 2019). The initial stage of PQDs recognition used signal processing techniques for extracting several characteristics/feature attributes. Likewise, feature selection and classification are also important stages of PQDs classification (Jamali et al., 2018). In earlier studies, several articles have reported the importance of FS and the types of FS methods used for these particular applications (Mishra, 2019). Ibrahim and Morcos (2002) reported the performance and importance of different AI techniques for PQDs classification. On the other hand, the advancement of AI makes possible the development of DL architecture which is able to bypass FS methods.

Wang and Chen in Wang and Chen (2019) proposed a new method of PQDs detection and classification using DCNN where it bypasses the traditional feature extraction process based on signal processing techniques and FS method. The proposed detection system comprises of multiple units of DCNN framework where each unit comprises 1-D convolutional, pooling, and batch-normalization layers to attain multi-dimensional features and reduce over-fitting. The validation of the suggested method was carried out by considering the following challenging environments: (i) 16 kinds of single and multiple PQDs (ii) noisy environments (20 dB, 30 dB, 40 dB) (iii) PQDs on microgrid system (iv) real-time PQDs.

Shen et al. used Improved Principal Component Analysis (IPCA) and 1-Dimensional-CNN (1-D-CNN) for recognition of PQ disturbances (Shen et al., 2019). Here, the IPCA and 1-D CNN were used to extract several statistical features. Afterwards, 1-D CNN is further used to classify the 12 categories of single and multiple PIDs. The performance of the proposed approach was compared with SVM and other existing ML techniques whose output shows that the proposed approach outperforms to other ML-based approaches. Ma et al. (2017) used Stacked auto-encoder (SAE), as a DL structure for the extraction of the accurate and large dimensional feature vector of PQ disturbance automatic recognitions. By this approach, the requirement of optimal FS step in automatic fault classification can be shunned. In addition to SAE, the PSO method was used to assist the classification task. In Liu et al. (2018c), Liu and his co-authors suggested a hybrid approach based on singular spectrum analysis (SSA), curvelet transform (CT) and DCNNs. SSA is a non-parametric method and can be used for recognizing week transient PQDs. Initially, PQDs were decomposed using SSA and CT techniques. First, six-level coefficients using SSA and three-level coefficients using CT were designated at features vector in the subsequent step. Finally, these features were used as input to DCNN for the automatic classification of PQ disturbances. Authors in Mohan et al. (2017) studied the potential of DL architecture in recognizing the PQDs in smart-grid, as it has the inherent auto-learning capability of optimal features from raw input data. Therefore, several DL architectures were implemented in the paper such as CNN, RNN, I-RNN, LSTM, GRU, and CNN–LSTM. Similarly, Rajiv and Tripathi

(2019) presented a LSTM–CNN based hybrid method for automatic PQ classification and detection. Table 5 shows the summary of some selected DL application in PQ disturbance detection and classification approaches.

6. Microgrid

A microgrid is an emerging revolution in the EUI by which several shortcomings in old conventional-grid have been overcome. The microgrid concept is rapidly accepted and adopted all over the world. Microgrid concept allows multiple numbers and types of energy resources integration to main-grid. Therefore, it is highly necessary to address the control and protection challenges and remedial actions. Moreover, the existence of multiple stakeholders in a microgrid is also a challenging task for proper energy management in a microgrid. AI applications against these challenges have proven to be one of the smartest approaches in recent times. Several articles have reported the applicability and advantages of AI in microgrid control, protection, and management (May et al., 2018; Kantamneni et al., 2015; Garduno-Ramirez and Borunda, 2017). Deep learning is one of the most advanced versions of AI which needs to be explored more deeply for solving these particular challenges in microgrid environments. Although the scope of DL application in these domains is not fully explored yet, the authors have tried to frame the outcomes of accomplished microgrid challenges using DL in the following sections.

6.1. Microgrid energy management

Energy management is an important issue in microgrid (MG) environment. Venayagamoorthy et al. (2016) proposed online dynamic energy management system (DEMS) which comprises evolutionary adaptive dynamic programming with a reinforcement learning support, in order to realize optimal or near-optimal DEMS in both grid-connected and islanded mode of microgrid operation.

6.2. Planning and operation of energy storage in microgrids

Authors in Shekhar et al. (2017) and Mbuwir et al. (2017) applied reinforcement learning for regulating an energy storage device (ESD) in an MG to maximize the self-utilization of the electricity generation from the local PVGS and to minimize the electricity-tariff and enslavement on the main-grid. In a few cases, multiple numbers of ESDs with different operating characteristics are employed as a group in a microgrid to increase its flexibility. Qiu et al. (2015) proposed a novel approach based on reinforcement learning for automatic control of charging and discharging periods of the different ESDs to increase system efficacy. Keerthisinghe et al. (2018) presented an energy management approach for solar photovoltaic storage systems using a policy function approximation (PFA) algorithm using ML. In the ML framework; the RNN, DNN, ELM, ANN, and SVM were used and compared. A PV based microgrid system comprises both short and long-term storage facilities, but due to different uncertainty challenges like lack of knowledge about future electricity consumption, highly weather dependency and levels of irradiance the effective operation of storage devices is highly hampered. François-Lavet et al. addressed this issue and suggested one novel strategy based on RL (François-Lavet et al., 2016). The method was empirically proved in the case of a domestic building situated in Belgium. In Hua et al. (2019), authors have studied and proposed a novel strategy of the energy management solution for a typical scenario of energy internet (EI). The considered EI scenario has multiple interconnected sub-grids where each individual sub-grid comprises of several electrical devices, such as photovoltaic system, wind turbine generators, microturbines, fuel cells, distributed energy resources, battery energy systems, loads and energy routers. Repository data from (https://dataport.cloud/) were used to design the power flow model from the photovoltaic system, wind, and load for this particular application. Based on the energy management principle of energy internet, the desired targets for optimal energy management were formulated as cost functions arithmetically. Then, the respected energy management issue was formulated as an optimal control problem (OCP). As the articulated problem was complex in nature, the traditional approaches, e.g., particle swarm optimization, genetic algorithm, simulate anneal arithmetic, etc., seem to be inappropriate, the author applied the deep-RL method to resolve the OCP.

6.3. Energy trading

It is well accepted that the future energy-grid will have numerous MGs and therefore, the coordination between each other is highly essential with proper energy trading strategy. Wang et al. (2016c) presented a novel energy trading strategy using RL for each smart-MG. The main purpose of the stated research work was to make each microgrid become smarter by adding additional feature of self-choosing ability by which each smart-MG can able to decide a strategy with possibility to trade the energy in an autonomous market. The proposed approach is definitely able capitalize its average revenue.

Normally, in MG case the service provider (vendor) is accountable for buying electrical power from the utility enterprises and selling it to the consumer. However, in several times several challenges have arisen for both of them (vender and the clients') as a result of partial/inadequate info acquisition system and the different sorts of uncertainties in the MG. A novel solution (RL-based dynamic pricing algorithm) to this particular issue was addressed by Kim et al. in Kim et al. (2015), by which the need of priori info about MG can ignored during guiding the strategy to both vendor and customers. Price of electrical energy is the fundamental component in electricity market. The trading in electricity market is generally based on electricity price. The electricity price is attuned by deviation in supply and demand response. Moreover, for the EUI, the bidding strategy decides the level of profit, and therefore, prediction of most accurate bidding price is an important aspect in the process. This helps the electrical generating company by reducing the transaction risk and providing more opportunity in the market. Kuo and Huang in Kuo and Huang (2018) presented a strategy based on two DL framework; CNN and LSTM for electricity price forecasting. The results of the stated approach was compared with several other ML techniques such as ANN, DT, MLP, SVM, etc. and found to be more effective than others. Lago et al. (2018) studied and compared twenty seven approaches comprising both ML and DL based approaches for electricity price forecasting. The outcome of the study reveals that the DNN based approach outperforms compared to other contemporary methods.

6.4. Islanding detection

Islanding is a situation in microgrid architecture framework where the DERs and its' connected loads are being disconnected from main-grid. This situation may be intentional or un-intentional. The un-intentional situation/islanding condition is always creates a dangerous environment for utility workers and the electrical apparatus. Therefore, detection of islanding with a stipulated time (for example, 2 s as per IEEE standard) is highly essential for making smart-MGs. Several islanding detection approaches with different detection principles were suggested by researches in recent past (Mishra et al., 2019). Out of these principles, the application of computational intelligence techniques was found to be more effective one (Laghari et al., 2014). Kong et al. (2018) used auto-encoder and stacked auto-encoder based DL model for islanding detection approach as novel application in this issue. Until that time, a fresh feature extraction process based on the combination of multi-resolution singular spectrum entropy and wavelet-transform was proposed and implemented to extract the features vector. A comparative analysis between the proposed DL based method and other ML approaches such as DT and SVM was carried

Table 5
Summary of some selected DL application in PQ disturbance detection and classification approaches.

DL model (Proposed)	Type of PQ Signals	Number of class	Data processing unit	Benchmark models	Results		Remarks	Ref.
					Accuracy in % (Benchmark Models)	Accuracy in % (Proposed model)		
Deep-CNN	Synthetic and Real time	16	X	SAE LSTM	SAE: 99.14 (pure), 99.84 (40 dB), 98.82 (20 dB) LSTM: 99.83 (pure), 92.39 (40 dB) 96.94 (20 dB)	99.96 (pure), 99.95 (40 dB), 98.13 (20 dB)	• The test results using both synthetic and real time in microgrid environment proves its the effectiveness, and shown to be a prevailing method compared to other DL based methods and other ML based methods. • Results shows that the computational time of the suggested approach was pointedly smaller than the other traditional methods owing to its parallel computation skill of GPU and combination of three conventional steps.	Wang and Chen (2019)
1-D-CNN	Simulated	12	IPCA	IPCA-SVM	99.05 (pure), 98.87 (50 dB) 96.76 (20 dB)	99.92 (pure), 92.85 (50 dB) 96.76 (20 dB)	• Comparative results shows that it outperform to other signal processing and ML based method. • The computational time involve with this method is 0.432 s, where the SVM based method take 0.792 s	(Shen et al., 2019)
DCNN	Synthetic	31	SSA, CT	SVM	99.91 (for single disturbance) 98.62 (for multiple disturbances)	100 (for single disturbance) 98.52 (for multiple disturbances)	• The suggested approach is also able to detect PQ disturbances for extremely noisy conditions. • Although the authors stated that the DCNNs based approach have reduced computational complexity and increased learning ability, the article have lack of evidence to prove it.	Liu et al. (2018c)
SAE	Synthetic	7	X	NA	NA	99.75 (pure), 99.60 (30 dB) 98.52 (20 dB)	• No benchmark model was tested on the similar dataset to compare with the proposed model; however they have used some previously published articles for this task. • Only seven cases of PQ events were tested.	Ma et al. (2017)
LSTM–CNN hybrid model	Synthetic	5	X	LSTM CNN	LSTM: 95.60 (Pure) CNN: 97.00 (Pure)	98.90 (Pure)	• The method did not consider noisy data for testing • Only five type single PQ disturbances were tested.	Rajiv and Tripathi (2019)
CNN (ResNet)	Simulated	54	X	STFT ST DWT	STFT: 93.60 (30 dB) 93.68.00 (40 dB)	CNN 98.52 (30 dB) 98.45 (40 dB)	• No benchmark classification method is shown in this work. • The paper is tested on a wide range of single and multiple disturbance classification.	Gong and Ruan (2020)
CNN–LSTM	Both Synthetic and Real-time	11	X	CNN RNN LSTM GRU	98 (CNN), 91.5 (RNN); 96.7 (LSTM), 96.4 (GRU)	98.4 (CNN–LSTM)	• Comparative analysis with other ML approaches: NA • The CNN–LSTM approach outperform to other DL approach with loss of classification is only 15% • Accuracy is noted to be 91.9% for the testing of 3 class of real-time disturbances	Mohan et al. (2017)

out where the proposed approach shows better performance in both accuracy and detection time. Authors in Abdelsalam et al. (2020) have proposed the islanding detection approach based on discrete Fourier transform and LSTM. The performance of the suggested approach was compared with other baseline classification approach (SVM, DT and ANN) applied to similar feature vector. From the comparative analysis the paper concludes that stated approach was proficient and effective in identifying the islanding events with high accuracy and dependability with smaller detection time.

6.5. Fault detection, classification and location

Intelligent fault detection, classification and location are highly necessary for smart-MG for its effective control and operation (Patnaik et al., 2020). The integration of inverter-based DERs in MGs makes customary fault recognition systems unsuitable owing to their dependency on significant fault current levels. Several techniques were presented to resolve these issues and have been an emerging topic for research over last decade. With the hypotheses related with smart-MG appealing maturing alarm amongst EUI scientists, the significance of developing an smart fault monitoring and identification scheme capable of recognizing and locating different sorts of faulty events cannot be

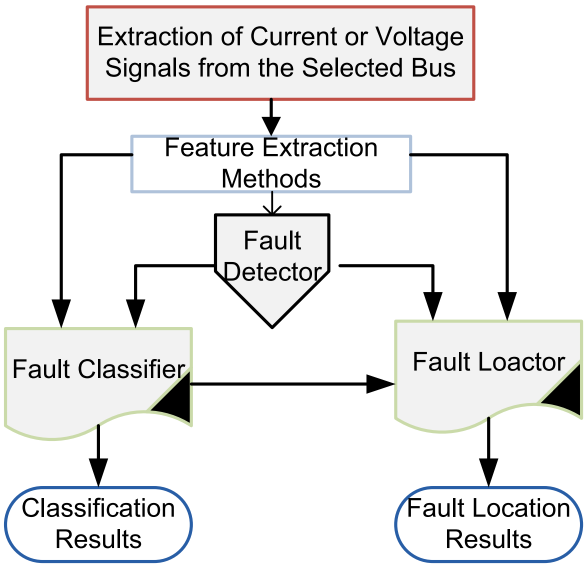


Fig. 7. Typical flowchart of fault detection, classification and location (Beheshtaein et al., 2019).

overstated. A basic framework for fault detection, classification and location is shown in Fig. 7. Several fault detection and classification approaches based on intelligent classifiers were reported in recent years (Beheshtaein et al., 2019; Hare et al., 2016; Chen et al., 2016). However, a few researches were reported in last couple of years where DL application was highlighted.

James et al. (2017) presented an intelligent fault detection approach for microgrid using discrete wavelet transform and DNN. The DNN used for this approach was mainly used GRU for its construction. The stated approach was able to handle three sub-problems such as fault type detection, faulty phase detection, and fault location. Lastly, the test results were also compared with DT, KNN, SVM, NBC, RF, Differential relay and Over-current (OC) relays, where the comparison result indicates the effectiveness of proposed DNN based approach in all three sub-problems. Similarly, the auto-encoder based deep learning neural network framework was proposed by Wang et al. (2016d) for power system fault diagnoses.

In the case of a PV based MG system, the V–I profile as a result of PV array fault and symmetrical line fault have high similarity which makes the protection scheme become more challenging. The traditional OC relays are unable to detect and classify such kinds of faults, and hence fail to offer discrete tripling signals for each event. To address this issue, Manohar et al. (2018) investigated and proposed a scheme using SAE and DNN which can differentiate PV array faults and symmetrical faults, and also achieve mode identification, fault recognition, categorization, and section identification issues. A similar research group (Manohar et al., 2020) in (2020) addressed another challenging issue "effect of irradiance variation on protection system" in a PV based MG system. In islanding condition, the irradiance variation pointedly disturbs the current contour in both robust and faulty situation, thus adding the risk of mal-operation of conventional OC relays. Therefore, a resilient and dependable MG mandates the development of a protection strategy, which is robust to changing irradiance levels. The use of CNN allows recognizing prejudiced features as of complicated datasets with decreased computational rate. Guo et al. (2017) proposed a new efficient technique of faulty feeder recognition in resonant grounding distribution systems (RGDS) using CWT and CNN. A brief study of recent use of the DL approach in fault detection, classification and location in different application domain of power system is presented in Table 6.

7. Electric vehicle

In smart-grid architecture, demand-side management (DSM) aims at suggesting inducement-based actions to change energy consumption patterns, such that more efficient utilization of energy is made. This repeatedly ensures that the current infrastructure is effectively used, to meet boosted demand and to minimalize investments in supplementary generation capacity (Zazo et al., 2016). However, the increasing use of Electric Vehicles (EVs) could enhance the possibility of overloading the power-grid by swelling demand peaks. Therefore, the necessity of DSM integration to EV charging becomes very high for supporting energy demand moderating, becoming the grid more economical, well-organized, and reliable.

López et al. (2018) proposed a smart charging scheme based on ML models for determining the charging sessions/time to charge the EV during connection sessions. This can be achieved by taking practical charging decisions through gaining the knowledge of several additional information (such as, historical data of connection sessions, environment, pricing, and demand time series, etc.), to decrease the total EV energy price. The initial stage of the scheme used dynamic programming to calculate the optimal solution of the historical connection sessions. Afterwards, ML models (such as DNN, Shallow-NN, and KNN) were trained to learn the patterns by which it can able to provide the right decision in real-time. Comparative results pointed towards DNN as a better performer compared to other ML models.

DSM of charging and discharging of electric vehicle encounters huge challenges due to the incomplete info about charging flexibility of EVs, plug-in times, power limitations, battery size, power curve, etc.. With an intention to overcome these issues, authors in Vandael et al. (2015) have proposed a scheme based on RL to recognize the EVs charging activities and then state an economic day-ahead consumption strategy using the learned behaviour.

8. Additional electrical power domain with DL application

This section reviewed several other research domain excluding the topics mentioned in Sections 3 to 7. This analysis or survey will clearly highlight the diversified scope of the deep learning/deep reinforcement learning applications in the electrical power domain.

8.1. Interval state estimation (ISE)

It is well known that the traditional power grid is being revolutionized on the way to smart-grid that allows mutual interactions amongst consumer and utility providers, and therefore more exposed to cyber-attacks. But, owing to the attacking cost, the attack approach might differ greatly from one operating scenario to another from the viewpoint of the adversary, which was not well-thought-out in preceding reports. Thus, Wang et al. (2018c) proposed a situation based two-stage sparse cyber-attack simulations for smart-grid with comprehensive network information. Afterwards, an effective interval state estimation (ISE) using a Sacked auto-encoder based defence mechanism was developed that helps to identify the probability for data manipulating.

8.2. Demand response

Recognizing and forecasting energy flexibility on the demand side has been important for applying demand response (DR). In the meantime, the development and implementation of SMs provide the consents to monitor the real-time power consumption level using Nonintrusive Load Monitoring (NILM). The NILM helps to determine the appliances are being operated and their specific consumption. Mocanu et al. (2016b) suggested Factored Four-Way- CRBMs to detect and forecast demand flexibility in the real-world. Tornai et al. (2017) used RNN to classify customers and shows that the stated systems were outperformed to other existing approaches; in few cases, the new considered

Table 6
DL approach for fault detection, classification and location in different application domain of power system.

Application	DL approach for fault classification	Brief description of process	Remarks	Ref.
Microgrid fault detection	Gated Recurrent Unit (GRU) based Deep Neural Network (DNN)	In this method, the extracted faulty feeders' currents are pre-processed by DWT to extract the statistical features. Afterwards, these features are fed as input to DNN to resister the fault information.	After a series of test simulations on 34 bus IEEE test system, it concludes that the stated approach is able to classify the fault types and detect the fault location with much better accuracy than other comparative methods.	Yu et al. (2019)
Power system fault Diagnosis	Stacked auto-encoders (SAE)	After the extraction of faulty data from SCADA centre these are processed by SAE based DNN for training and testing.	This method is aimed in such a way to achieve the solution to three major issues that are usually encounter by traditional NN approach such as availability of data, better local optimum, and diffusion of gradients.	Wang et al. (2016d)
Detection of array faults and symmetrical line faults in a PV integrated microgrid	Sparse auto encoder (Sparse-AE)	In the Initial phase, the measured voltage and current signals from the pre-defined relays are converted to grey-scale images. Afterwards, these images are processed through Sparse-AE to achieve unsupervised feature learning. By this process the method is able to discriminate the array faults and symmetrical line faults as well as accomplished the fault classification and section identification task.	The method can help to improve the reliability of the fault protection scheme in a SPV based microgrid. The comparative performance analyses with other ML approaches such as ANN, SVM, and DT in both the mode of microgrid operation prove the superiority of the stated approach.	Manohar et al. (2018)
Earth Fault Detection in RGDS	Convolutional Neural Network (CNN)	The multi-resolution based grey-scale images were extracted from transient zero-sequence current signals of the faulty and healthy feeder using DWT. Afterwards; CNN was applied for feature extraction from the extracted greyscale images and subsequently, the classification of the healthy and faulted feeder.	The simulation results justify the reliability, robustness and accuracy of stated fault detection approach in diverse operating conditions, such as two-point grounding fault, different network structures, transformer reversed, arc grounding fault, etc. Moreover, the comparative performance analysis based on traditional ML algorithm, such as the Adaboost or SVM were also conducted to prove the effectiveness of the proposed approach.	Guo et al. (2018)
Protection scheme for PV integrated microgrid under solar irradiance intermittency	Convolutional neural network (ConvNet)	In the Initial phase, the measured voltage and current signals from the pre-defined relays are converted to grey-scale images and the patches so obtained from the multichannel training dataset containing spatial and temporal information. Afterwards, these dataset are processed through ConvNet to achieve unsupervised feature learning. By this process the method is able to achieve the fault classification and section identification task in islanded microgrid.	The OPAL-RT simulation environment has been used to validate the accuracy of the proposed approach. The overall accuracy obtained through the proposed approach is 99.48% and 99.58% in fault detection and section identification task respectively.	Manohar et al. (2020)
Intelligent microgrids protection	Deep feed forward neural network (DFFNN) CNN based DNN	Initially, the three-phase current signals are extracted at relaying end and feed to the proposed algorithm. Afterwards, these information are accesses by the DNN assigned for fault detection task. As soon as the fault is detected by the algorithm, the fault type information is registered for further use.	Obtained results show a encouraging performance regardless a wide variations in fault resistance and DER penetrations compared to existing approaches. Furthermore, the detection time was found to be less than 2 cycles from the point of fault inception.	Samal et al. (2019)

rates can reach closely 100%. In the point of fact, DR has belonged to high dimensional control problems which may encounter the challenges of restricted observability and uncertainty. To control such types of challenges, several authors have suggested different deep reinforcement learning (DRL) based approaches to set the residential DR at the apparatus level Wen et al., 2015; Ruelens et al., 2014; Costanzo et al., 2016. In the DRL approach, the deep learning architecture is combined with a reinforcement learning algorithm (such as, Q-learning and Markov Decision Process). Mocanu et al. (2018) suggested deep Q-learning (DQL) and Deep policy gradient (DPG) for decision making policies at both the individual residents and the aggregated level. The outcome of the study concludes that the DPG outperforms DQL in live scheduling of DERs at both the levels.

8.3. Internet of things and Cognitive networks

Cognitive networks (CNs) play a key role in the Internet of Things (IoT) architecture applications, such as healthcare management, agricultural science, ecosystem monitoring, and smart-metring. On the other hand, the present small packet transmission efficacy of the Internet of Things encounters a challenge of the overcrowded spectrum for the fast-growing reputations of numerous wireless appliances. Nevertheless, the development of a hybrid approach i.e. cognitive radio-based IoT (CIoT) technology is becoming an encouraging solution to the above-mentioned issue. Again, to increase the packet transmission efficacy utilizing CNs is a foremost challenge for CIoT. In this regard, Zhu et al. (2017) proposed a novel Q-learning-based communication scheduling procedure employing the DL technique for the CIoT to resolve the difficulty of achieving the right strategy for packets transmission of different buffers via numerous networks to increase the system output.

8.4. Fuel cell fault diagnosis

The development of fault diagnosis tools is very important for confirming and increasing the stability, durability, and permanency of solid oxide fuel cell (SOFC) systems. Zhang et al. (2019b) proposed a novel method based on DL for diagnosing the presence of simultaneous faults. The sparse auto-encoder (SSAE) was used for the automatic extraction of unique features. Afterwards, the K-binary classifier is used for detecting simultaneous faults.

In point of fact, the fuel cells are invincible to the impurities of hydrogen and operational environments. Due to this, the output performance of fuel cells is highly affected and degraded with moving time. Therefore, the prediction of performance degradation is becoming a challenging task for the researcher to increase the reliability of fuel cells. Authors in Ma et al. (2018) proposed an advanced fuel cell degradation forecasting scheme based on Grid-LSTM–RNN. The comparative result shows that the stated scheme was able to predict the degradation rate for a longer duration.

8.5. Anomaly detection and fault analysis of wind turbine

The insensitive and inconsistent operating environment of wind turbines leads to malfunctions of connected apparatuses, such as the gearbox, main bearing, generator, inverter, and controller (Kusiak and Verma, 2011; Zhao and Li, 2017). Therefore, the uninterrupted health-care monitoring of the wind turbine system via premature fault detection approaches might help to enhance the reliability and decrease maintenance charges of the turbine prior to they face a catastrophic point. In this regard, Zhao et al. (2018) suggested a DL approach based on deep-auto-encoder (DAE) for anomaly and fault detection schemes for the wind turbine. The required data from the wind turbines were collected through operational supervisory control and data acquisition (SCADA) systems. Authors in Cheng et al. (2017) proposed an innovative fault diagnosis technique for the drivetrain gearboxes of the wind turbine connected with doubly-fed induction generators (DFIGs) using a rotor current signal (RCS) analysis. The stated work was accomplished by following steps; (i) extraction of the instantaneous fundamental frequency of the RCS (ii) application of Hilbert transform for demodulation of RCS to obtains its envelope containing fault characteristic frequencies (iii) power spectral density analysis was performed on the resampled envelope signal for the gearbox fault detection (iv) application of DL architecture based classifier that consists of an SAE and an SVM for gearbox fault classification using obtained fault features.

9. Critical discussion

This section highlights some key facts about the DL applications in EUI through statistical analyses like distribution of articles by different DL architecture versus EUI application domain. This will provide a clearer idea to the researcher/scientist regarding the application feasibility or scope of application and the acknowledgement received from the popular publishing houses. In addition to this, several open research issues are pointedly explained that may help the researcher in their future research problem formulation.

9.1. Distribution of articles by DL architecture versus EUI application domain

The distribution of articles based on DL based neural network architectures in different application domain in EUI is offered in Table 7. It can be clearly observed from the table that the advanced forecasting support to renewable power generation has been most popular research domain for DL architecture. Moreover, it can be also analysed that the LSTM and CNN models were most frequently used DL based model in EUI.

9.2. Open research issues and further discussion

This article has presented an organized review for burgeoning DL based approaches applied to the electrical utility industry. It can be concluded that there has been lots of interest in using DL architectures for delivering solutions to the problems associated with the electrical power domain. The CNN, RNN, autoencoders, DBN and deep Q-learning are the most commonly used approaches in the DL framework. Even though several constructive outcomes are documented from the above-mentioned reviewed stuff, a few additional concerns regarding DL applications are highlighted below:

- From the above study, it can be clearly analysed that each and every method in DL framework are viable candidates for providing more efficient performance compared to conventional ML approaches for both prediction and classification problems. Deep learning systems are able to handle large volumes of dataset extracted from SMs and other data sources. But, their performances require to be further examined using more complex dataset. Moreover, the question regarding the selecting criteria for the number of layers and parameters are yet to be answered.
- In the forecasting support sections, it can be analysed that most of the work till date have concentrated on short-time time-horizons and univariate time series prediction. The development of medium-term and long-term forecasting methods with negligible error, and the multivariate time-series prediction are few challenging tasks which need to be addressed through DL application in near future.
- Demand response is belongs to high dimensional control problems which may encounters the challenges of partial observability and uncertainty. A few research based on DRL has already proposed by different authors, but how to analyse flexibility as of the consumer viewpoints, direct control of load, and accurate decision making for each stakeholder in DR, specifically in the partial observability situations, are still need to be addressed.
- As per the reviews carried out in Sections 6.4, 6.5 and 8.5, DL has been proved to be an effective tool for fault/anomaly detection and classification. But, a few concerns that need to be answered such as small sample learning problems, detection of differences between normal and pre-fault condition and feasibility of real-time applications.
- It has been seen that the model-free methods such as RL and DRL are proven to be most effective and alternative to model based methods. However, the deep learning and reinforcement learning methods have the characteristics of non-interpretability, and therefore, reliability on such black box approaches for EPS control and analysis is still being debatable.
- The application of DL in the field of image processing and speech processing is highly successful, but as compared to these fields the complexity in power system is found to be slightly higher. This is for the reason that the characteristics of fault currents/voltages can be heterogeneous and vary along with different faulty situations and time period. Therefore, the black box solution is necessary for these types of problems which necessitate lots of research effort from the researches belong to industry and academic background.
- From the above-mentioned review, it is also concluded that the research emphasis on cost of implementation, reasons behind selection of DL architectures for the specific issue, and reducing the computation complexity is highly required.
- Moreover, It is well knows that the implementation time and fault detection response time are equally important in several problem belongs to electrical domain. Therefore, there is an essential to pay attention to the structure of the computation to minimize time complexity.

Table 7
Status of the DL model applied to various research fields under the area of electrical domains.

Application domain	Ref.	AEs	CNNs/DCNN	DNN	Deep-RNN	LSTM	GRU	RBM	DBN	DBM	DQL/DRL
Load forecasting	Guo et al. (2018)			√							
	Hossen et al. (2017)			√							
	Din and Marnerides (2017)			√							
	He (2017)		√								
	He et al. (2017)								√		
	(Dedine et al. 2016)								√		
	Wen et al. (2020)						√				
	Bedi and Toshniwal (2018)					√					
	Qiu et al. (2017)								√		
	Estebsari and Rajabi (2020)					√	√				
	Tong et al. (2018)	√									
	Marino et al. (2016)					√					
	Amarasinghe et al. (2017)		√			√					
	Kong et al. (2019)					√					
	Mnih et al. (2012)							√			
	Taylor et al. (2011)							√			
	Mocanu et al. (2016a)							√			
	Shi et al. (2017)				√						
	Bouktif et al. (2018)					√					
	Kumar et al. (2018)						√				
Wind power forecasting	Wang et al. (2016b)								√		
	Liu et al. (2018a)					√					
	Liu et al. (2018b)		√								
	Chen et al. (2018)					√					
	Zhu et al. (2018)		√								
	Feng et al. (2017)										
	Torres et al. (2018)			√							
	Yu et al. (2019)					√					
	Wu et al. (2016)				√	√					
	Wang et al. (2017a)		√								
	Zhou et al. (2018)		√								
	Wang et al. (2017a)		√								
	Wang et al. (2018b)								√		
	Qureshi et al. (2017)								√		
	Sergio and Ludermir (2015)		√						√		
	Hu et al. (2016)		√								
	Khodayar et al. (2017)		√								
	Huang et al. (2018)					√					
	Zhang et al. (2015)									√	
	Hu and Chen (2018)					√					
	Memarzadeh and Keynia (2020)					√					
	Liu et al. (2020)		√			√					
	Peng et al. (2020)					√	√				
	Hu et al. (2020)										
	Ma et al. (2020)					√					
	Mi and Zhao (2020)					√					
	Lin et al. (2020)			√							
	Jiajun et al. (2020)								√		
	Jahangir et al. (2020)	√									
	Devi et al. (2020)					√					
PV power forecasting	Abdel-Nasser and Mahmoud (2019)				√	√					
	Gensler et al. (2016)				√	√			√		
	Wen et al. (2019)										
	Huang and Kuo (2019)			√							
	Li et al. (2017)								√		
	Wang et al. (2017b)		√								
	Zang et al. (2018)			√							
	Zhen et al. (2020)		√			√					
	Alzahrani et al. (2017)				√						
	Aprillia et al. (2020)		√			√					
	Aslam et al. (2020)					√	√				
	Lima et al. (2020)					√					
	Hossain and Mahmood (2020)					√					
	Li et al. (2020)				√	√	√				
	Zang et al. (2020)		√								
Power system state forecasting	Zhang et al. (2019a)				√						
Power quality detection	Wang and Chen (2019)		√								
	(Shen et al., 2019)		√								
	Ma et al. (2017)	√									
	Liu et al. (2018c)		√								
	Mohan et al. (2017)		√		√	√	√				
	Rajiv and Tripathi (2019)										
	Gong and Ruan (2020)		√								

(*continued on next page*)

Table 7 (*continued*).

Application domain	Ref.	AEs	CNNs/DCNN	DNN	Deep-RNN	LSTM	GRU	RBM	DBN	DBM	DQL/DRL
Microgrid energy management	May et al. (2018)										√
Planning and operation of energy storage in microgrids	Keerthisinghe et al. (2018)			√	√						
	François-Lavet et al. (2016)										√
	Hua et al. (2019)										√
Energy trading	Kim et al. (2015)										√
	Kuo and Huang (2018)		√			√					
	Lago et al. (2018)			√							
Islanding detection	Kong et al. (2018)	√									
	Abdelsalam et al. (2020)					√					
Microgrid fault detection, classification and location	James et al. (2017)			√			√				
	Wang et al. (2016d)	√									
	Manohar et al. (2018)	√		√							
	Manohar et al. (2020)		√								
	Guo et al. (2017)		√								
Electric vehicle	López et al. (2018)			√							
	Vandael et al. (2015)										√
Interval state estimation	Wang et al. (2018c)	√									
Demand response	Tornai et al. (2017)				√						
	Wen et al. (2015)										√
	Ruelens et al. (2014)										√
	Costanzo et al. (2016)										√
	Mocanu et al. (2018)										√
Internet of Things and Cognitive networks	Zhu et al. (2017)										√
Fuel cell fault diagnosis	Zhang et al. (2019b)	√									
	Ma et al. (2018)				√	√					
Anomaly detection and fault analysis of wind turbine	Zhao et al. (2018)	√									
	Cheng et al. (2017)	√									

- It is well-known that the SG, as cyber–physical critical infrastructures claims better reliability and efficiency through effective use of AI. However, the dependency on AI has leads to several other issues like increasing vulnerability to malicious attacks or cyber-attack. In this regards, SG integrates the physical power systems and cyber systems, which exhibits the following characteristics like Self-adaption, self-organization, and self-learning. Due to this features, the cyber–physical can effectively deal with the faulty events, cyber-attacks, and power-crises, with the intention to increase the resiliency of SG infrastructure. In this direction, the ML approaches play an impotent role in the cyber–physical strategy to deal with the false data injection attack in SGs. Although several ML based cyber–physical strategy already suggested by various researchers in past, the development of DL based cyber–physical strategy is still an open research area.

For the electrical utility industry, DL has opened a plethora of futuristic research opportunities; the important questions centering around the impact of DL based innovation. The above-mentioned key challenges give way to a number of future research opportunities. With the continued research interest, the eventual goal is to develop an overall solution with several interacting components. The above mentioned key challenges show the directions towards numbers of forthcoming research openings.

It can be clearly analysed from Table 7 that, the DL methods such as autoencoder, DBN, CNN, and RNN has been employed for designing model to solve the diversified electrical domain problems. The researchers/authors have used these techniques as novel tools and attempt to solve the specific problem. Here, a question may arise that "should there be any major change in the outcome, irrespective of the chosen architecture?" These types of comparative analyses are missing in many of the reviewed articles.

10. Conclusion

Recently, deep learning methods have gained huge attention throughout the globe. In contrast to conventional machine learning (AI-1.0), deep learning/deep reinforcement learning (AI-2.0) techniques are more application-oriented, and therefore the adoption rate by the advanced manufacturer is increased in recent years. In this survey article, the authors have presented a widespread review of deep learning applications to several distinct problems associated with the electrical power domain. The work emphasized on the feasibility study and summarizes the recent developments in this area. In the DL application research domain, we have studied almost related areas under the umbrella of the electrical utility industry. From a technological perspective, these application domains are characterized as advanced forecasting supports which include wind/PV power predictions, anomaly detection, demand response, cybersecurity, microgrid management, fault detection in microgrid, electric vehicle and energy storage management. In addition, we have presented several key challenges which directed towards several open research scope. The development of deep learning architecture for the solution of electrical domain problems is still an emergent and encouraging research field. We hope that this survey can helpful for the readers/researchers to gain an absolute picture and deep vision into this area.

CRediT authorship contribution statement

Manohar Mishra: Writing - review & editing, Writing - original draft. **Janmenjoy Nayak:** Conceptualization, Writing - review & editing. **Bighnaraj Naik:** Methodology, Writing - original draft. **Ajith Abraham:** Supervision, Investigation, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abdel-Aal, R.E., Al-Garni, A.Z., 1997. Forecasting monthly electric energy consumption in eastern Saudi Arabia using univariate time-series analysis. Energy 22 (11), 1059–1069.

Abdel-Nasser, M., Mahmoud, K., 2019. Accurate photovoltaic power forecasting models using deep LSTM-RNN. Neural Comput. Appl. 31 (7), 2727–2740.
Abdelsalam, A.A., Salem, A.A., Oda, E.S., Eldesouky, A.A., 2020. Islanding detection of microgrid incorporating inverter based DGs using long short-term memory network. IEEE Access.
Ahmed, R., Sreeram, V., Mishra, Y., Arif, M.D., 2020. A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization. Renew. Sustain. Energy Rev. 124, 109792.
Akhter, M.N., Mekhilef, S., Mokhlis, H., Shah, N.M., 2019. Review on forecasting of photovoltaic power generation based on machine learning and metaheuristic techniques. IET Renew. Power Gener. 13 (7), 1009–1023.
Almalaq, A., Edwards, G., 2017. A review of deep learning methods applied on load forecasting. In: 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE, pp. 511–516.
Almalaq, A., Zhang, J.J., 2020. Deep learning application: Load forecasting in big data of smart grids. In: Deep Learning: Algorithms and Applications. Springer, Cham, pp. 103–128.
Alom, M.Z., Taha, T.M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M.S., Hasan, M., Van Essen, B.C., Awwal, A.A., Asari, V.K., 2019. A state-of-the-art survey on deep learning theory and architectures. Electronics 8 (3), 292.
Alzahrani, A., Shamsi, P., Dagli, C., Ferdowsi, M., 2017. Solar irradiance forecasting using deep neural networks. Procedia Comput. Sci. 114, 304–313.
Amarasinghe, K., Marino, D.L., Manic, M., 2017. Deep neural networks for energy load forecasting. In: 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE). IEEE, pp. 1483–1488.
Angelopoulos, M., Kontakou, C., Pollalis, Y., 2019. Digital Transformation and Lean Management. Challenges in the Energy Industry of Utilities. A Review (No. 95523). University Library of Munich, Germany.
Aprillia, H., Yang, H.T., Huang, C.M., 2020. Short-term photovoltaic power forecasting using a convolutional neural network–salp swarm algorithm. Energies 13 (8), 1879.
Aslam, M., Lee, J.M., Kim, H.S., Lee, S.J., Hong, S., 2020. Deep learning models for long-term solar radiation forecasting considering microgrid installation: A comparative study. Energies 13 (1), 147.
Baliyan, A., Gaurav, K., Mishra, S.K., 2015. A review of short term load forecasting using artificial neural network models. Procedia Comput. Sci. 48, 121–125.
Barakat, E.H., 2001. Modeling of nonstationary time-series data. Part II. Dynamic periodic trends. Electr. Power Energy Syst. 23 (1), 63–68.
Bedi, J., Toshniwal, D., 2018. Empirical mode decomposition based deep learning for electricity demand forecasting. IEEE Access 6, 49144–49156.
Beheshtaein, S., Cuzner, R., Savaghebi, M., Guerrero, J.M., 2019. Review on microgrids protection. IET Gener. Transm. Distrib. 13 (6), 743–759.
Bose, B.K., 2017. Artificial intelligence techniques in smart grid and renewable energy systems—Some example applications. Proc. IEEE 105 (11), 2262–2273.
Botvinick, M., Ritter, S., Wang, J.X., Kurth-Nelson, Z., Blundell, C., Hassabis, D., 2019. Reinforcement learning, fast and slow. Trends Cogn. Sci..
Bouktif, S., Fiaz, A., Ouni, A., Serhani, M., 2018. Optimal deep learning lstm model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches. Energies 11 (7), 1636.
Bouwmans, T., Javed, S., Sultana, M., Jung, S.K., 2019. Deep neural network concepts for background subtraction: A systematic review and comparative evaluation. Neural Netw..
Cavallaro, F., 2005. Electric load analysis using an artificial neural network. Int. J. Energy Res. 29 (5), 377–392.
CER, 2011. Electricity smart metering customer behaviour trials (CBT) findings report. In: Commission Energy Regulation. Tech. Rep. CER11080b, Commission Energy Regulation Dublin, Ireland.
Chen, Q., Folly, K.A., 2018. Wind power forecasting. IFAC-PapersOnLine 51 (28), 414–419.
Chen, K., Huang, C., He, J., 2016. Fault detection, classification and location for transmission lines and distribution systems: a review on the methods. High Volt. 1 (1), 25–33.
Chen, J., Zeng, G.Q., Zhou, W., Du, W., Lu, K.D., 2018. Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization. Energy Convers. Manage. 165, 681–695.
Cheng, F., Wang, J., Qu, L., Qiao, W., 2017. Rotor-current-based fault diagnosis for DFIG wind turbine drivetrain gearboxes using frequency analysis and a deep classifier. IEEE Trans. Ind. Appl. 54 (2), 1062–1071.
Costanzo, G.T., Iacovella, S., Ruelens, F., Leurs, T., Claessens, B.J., 2016. Experimental analysis of data-driven control for a building heating system. Sustain. Energy Grids Netw. 6, 81–90.
Das, U.K., Tey, K.S., Seyedmahmoudian, M., Mekhilef, S., Idris, M.Y.I., Van Deventer, W., Horan, B., Stojcevski, A., 2018. Forecasting of photovoltaic power generation and model optimization: A review. Renew. Sustain. Energy Rev. 81, 912–928.
Daut, M.A.M., Hassan, M.Y., Abdullah, H., Rahman, H.A., Abdullah, M.P., Hussin, F., 2017. Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review. Renew. Sustain. Energy Rev. 70, 1108–1118.
Dedinec, A., Filiposka, S., Dedinec, A., Kocarev, L., 2016. Deep belief network based electricity load forecasting: An analysis of Macedonian case. Energy 115, 1688–1700.
Deng, L., Liu, Y. (Eds.), 2018. Deep Learning in Natural Language Processing. Springer.
Devi, A.S., Maragatham, G., Boopathi, K., Rangaraj, A.G., 2020. Hourly day-ahead wind power forecasting with the EEMD-CSO-LSTM-EFG deep learning technique. Soft Comput. 1–21.
Dijk, J., Schutte, K., Oggero, S., 2019. A vision on hybrid AI for military applications. In: Artificial Intelligence and Machine Learning in Defense Applications, Vol. 11169. International Society for Optics and Photonics, 111690D.
Din, G.M.U., Marnerides, A.K., 2017. Short term power load forecasting using deep neural networks. In: 2017 International Conference on Computing, Networking and Communications (ICNC). IEEE, pp. 594–598.
Ding, L., Xiao, L., Zhou, K., Lan, Y., Zhang, Y., 2018. A new RNN model with a modified nonlinear activation function applied to complex-valued linear.
Dong, X., Qian, L., Huang, L., 2017. Short-term load forecasting in smart grid: A combined CNN and K-means clustering approach. In: 2017 IEEE International Conference on Big Data and Smart Computing (BigComp). IEEE, pp. 119–125.
Elsheikh, A.H., Sharshir, S.W., Elaziz, M.A., Kabeel, A.E., Guilan, W., Haiou, Z., 2019. Modeling of solar energy systems using artificial neural network: A comprehensive review. Sol. Energy 180, 622–639.
Estebsari, A., Rajabi, R., 2020. Single residential load forecasting using deep learning and image encoding techniques. Electronics 9 (1), 68.
Feng, C., Cui, M., Hodge, B.M., Zhang, J., 2017. A data-driven multi-model methodology with deep feature selection for short-term wind forecasting. Appl. Energy 190.
Filik, Ü.B., Gerek, Ö.N., Kurban, M., 2011. A novel modeling approach for hourly forecasting of long-term electric energy demand. Energy Convers. Manage. 52 (1), 199–211.
François-Lavet, V., Taralla, D., Ernst, D., Fonteneau, R., 2016. Deep reinforcement learning solutions for energy microgrids management. In: European Workshop on Reinforcement Learning (EWRL 2016).
Garduno-Ramirez, R., Borunda, M., 2017. An intelligent multi-agent based improved approach for conventional and renewable power generation operation and control. J. Renew. Sustain. Energy 9 (1), 015302.
Gensler, A., Henze, J., Sick, B., Raabe, N., 2016. Deep learning for solar power forecasting—An approach using autoencoder and LSTM neural networks. In: 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC).
Gong, R., Ruan, T., 2020. A new convolutional network structure for power quality disturbance identification and classification in micro-grids. IEEE Access.
Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT press.
Guo, M.F., Zeng, X.D., Chen, D.Y., Yang, N.C., 2017. Deep-learning-based earth fault detection using continuous wavelet transform and convolutional neural network in resonant grounding distribution systems. IEEE Sens. J. 18 (3), 1291–1300.
Guo, Z., Zhou, K., Zhang, X., Yang, S., 2018. A deep learning model for short-term power load and probability density forecasting. Energy 160, 1186–1200.
Haque, M.T., Kashtiban, A.M., 2000. Application of neural networks in power systems; A review. Power 2005.
Hare, J., Shi, X., Gupta, S., Bazzi, A., 2016. Fault diagnostics in smart micro-grids: A survey. Renew. Sustain. Energy Rev. 60, 1114–1124.
He, W., 2017. Load forecasting via deep neural networks. Procedia Comput. Sci. 122, 308–314.
He, Y., Deng, J., Li, H., 2017. Short-term power load forecasting with deep belief network and copula models. In: 2017 9th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), Vol. 1. IEEE, pp. 191–194.
Hossain, M.S., Mahmood, H., 2020. Short-term photovoltaic power forecasting using an LSTM neural network. In: 2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT). IEEE, pp. 1–5.
Hossen, T., Plathottam, S.J., Angamuthu, R.K., Ranganathan, P., Salehfar, H., 2017. Short-term load forecasting using deep neural networks (DNN). In: 2017 North American Power Symposium (NAPS). IEEE, pp. 1–6.
Hu, Y.L., Chen, L., 2018. A nonlinear hybrid wind speed forecasting model using LSTM network, hysteretic ELM and differential evolution algorithm. Energy Convers. Manage. 173, 123–142.
Hu, H., Wang, L., Lv, S.X., 2020. Forecasting energy consumption and wind power generation using deep echo state network. Renew. Energy.
Hu, Q., Zhang, R., Zhou, Y., 2016. Transfer learning for short-term wind speed prediction with deep neural networks. Renew. Energy 85, 83–95.
Hua, H., Qin, Y., Hao, C., Cao, J., 2019. Optimal energy management strategies for energy internet via deep reinforcement learning approach. Appl. Energy 239, 598–609.
Huang, C.J., Kuo, P.H., 2019. Multiple-input deep convolutional neural network model for short-term photovoltaic power forecasting. IEEE Access 7, 74822–74834.
Huang, Y., Liu, S., Yang, L., 2018. Wind speed forecasting method using EEMD and the combination forecasting method based on GPR and LSTM. Sustainability 10 (10), 3693.
Ibrahim, W.A., Morcos, M.M., 2002. Artificial intelligence and advanced mathematical tools for power quality applications: a survey. IEEE Trans. Power Deliv. 17 (2), 668–673.
Jager, D., Andreas, A., 1996. NREL National Wind Technology Center (NWTC): M2 Tower; Boulder, Colorado (Data). http://dx.doi.org/10.5439/1052222.
Jahangir, H., Golkar, M.A., Alhameli, F., Mazouz, A., Ahmadian, A., Elkamel, A., 2020. Short-term wind speed forecasting framework based on stacked denoising auto-encoders with rough ANN. Sustain. Energy Technol. Assess. 38, 100601.

Jamali, S., Farsa, A.R., Ghaffarzadeh, N., 2018. Identification of optimal features for fast and accurate classification of power quality disturbances. Measurement 116, 565–574.
James, J.Q., Hou, Y., Lam, A.Y., Li, V.O., 2017. Intelligent fault detection scheme for microgrids with wavelet-based deep neural networks. IEEE Trans. Smart Grid 10 (2), 1694–1703.
Jiajun, H., Chuanjin, Y., Yongle, L., Huoyue, X., 2020. Ultra-short term wind prediction with wavelet transform, deep belief network and ensemble learning. Energy Convers. Manage. 205, 112418.
Kalogirou, S.A., 2001. Artificial neural networks in renewable energy systems applications: a review. Renew. Sustain. Energy Rev. 5 (4), 373–401.
Kang, J., Zhao, H., 2012. Application of improved grey model in long-term load forecasting of power engineering. Syst. Eng. Proc. 3, 85–91.
Kantamneni, A., Brown, L.E., Parker, G., Weaver, W.W., 2015. Survey of multi-agent systems for microgrid control. Eng. Appl. Artif. Intell. 45, 192–203.
Keerthisinghe, C., Chapman, A.C., Verbič, G., 2018. Energy management of PV-storage systems: Policy approximations using machine learning. IEEE Trans. Ind. Inf. 15 (1), 257–265.
Khan, S., Yairi, T., 2018. A review on the application of deep learning in system health management. Mech. Syst. Signal Process. 107, 241–265.
Khatoon, S., Singh, A.K., 2014. Effects of various factors on electric load forecasting: An overview. In: 2014 6th IEEE.
Khodayar, M., Kaynak, O., Khodayar, M.E., 2017. Rough deep neural architecture for short-term wind speed forecasting. IEEE Trans. Ind. Inf. 13 (6), 2770–2779.
Khuntia, S.R., Rueda, J.L., van der Meijden, M.A., 2016. Forecasting the load of electrical power systems in mid-and long-term horizons: a review. IET Gener. Transm. Distrib. 10 (16), 3971–3977.
Kim, B.G., Zhang, Y., Van Der Schaar, M., Lee, J.W., 2015. Dynamic pricing and energy consumption scheduling with reinforcement learning. IEEE Trans. Smart Grid 7 (5), 2187–2198.
Kong, W., Dong, Z.Y., Jia, Y., Hill, D.J., Xu, Y., Zhang, Y., 2019. Short-term residential load forecasting based on LSTM recurrent neural network. IEEE Trans. Smart Grid 10 (1), 841–851.
Kong, X., Xu, X., Yan, Z., Chen, S., Yang, H., Han, D., 2018. Deep learning hybrid method for islanding detection in distributed generation. Appl. Energy 210, 776–785.
Kumar, S., Hussain, L., Banarjee, S., Reza, M., 2018. Energy load forecasting using deep learning approach-LSTM and GRU in spark cluster. In: 2018 Fifth International Conference on Emerging Applications of Information Technology (EAIT). IEEE, pp. 1–4.
Kumar, V.K., Srivastava, S.C.D., 1999. Power system state forecasting using artificial neural networks. Electr. Mach. Power Syst. 27 (6), 653–664.
Kuo, P.H., Huang, C.J., 2018. An electricity price forecasting model by hybrid structured deep neural networks. Sustainability 10 (4), 1280.
Kusiak, A., Verma, A., 2011. A data-mining approach to monitoring wind turbines. IEEE Trans. Sustain. Energy 3 (1), 150–157.
Laghari, J.A., Mokhlis, H., Karimi, M., Bakar, A.H.A., Mohamad, H., 2014. Computational intelligence based techniques for islanding detection of distributed generation in distribution network: A review. Energy Convers. Manage. 88, 139–152.
Lago, J., De Ridder, F., De Schutter, B., 2018. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. Appl. Energy 221, 386–405.
LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. nature 521 (7553), 436–444.
Lee, D.G., Lee, B.W., Chang, S.H., 1997. Genetic programming model for long-term forecasting of electric power demand. Electr. Power Syst. Res. 40 (1), 17–22.
Lei, M., Shiyan, L., Chuanwen, J., Hongling, L., Yan, Z., 2009. A review on the forecasting of wind speed and generated power. Renew. Sustain. Energy Rev. 13 (4), 915–920.
Li, L.L., Cheng, P., Lin, H.C., Dong, H., 2017. Short-term output power forecasting of photovoltaic systems based on the deep belief net. Adv. Mech. Eng. 9 (9), 1687814017715983.
Li, P., Zhou, K., Lu, X., Yang, S., 2020. A hybrid deep learning model for short-term PV power forecasting. Appl. Energy 259, 114216.
Lima, M.A.F., Carvalho, P.C., Fernández-Ramírez, L.M., Braga, A.P., 2020. Improving solar forecasting using deep learning and portfolio theory integration. Energy 195, 117016.
Lin, Z., Liu, X., Collu, M., 2020. Wind power prediction based on high-frequency SCADA data along with isolation forest and deep learning neural networks. Int. J. Electr. Power Energy Syst. 118.
Liu, H., Chen, C., 2019. Data processing strategies in wind energy forecasting models and applications: A comprehensive review. Appl. Energy 249, 392–408.
Liu, H., Chen, C., Lv, X., Wu, X., Liu, M., 2019. Deterministic wind energy forecasting: A review of intelligent predictors and auxiliary methods. Energy Convers. Manage. 195, 328–345.
Liu, H., Hussain, F., Shen, Y., Arif, S., Nazir, A., Abubakar, M., 2018c. Complex power quality disturbances classification via curvelet transform and deep learning. Electr. Power Syst. Res. 163, 1–9.
Liu, H., Mi, X.W., Li, Y.F., 2018a. Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network. Energy Convers. Manage. 156, 498–514.
Liu, H., Mi, X., Li, Y., 2018b. Smart deep learning based wind speed prediction model using wavelet packet decomposition, convolutional neural network and convolutional long short term memory network. Energy Convers. Manage. 166, 120–131.
Liu, Y., Qin, H., Zhang, Z., Pei, S., Jiang, Z., Feng, Z., Zhou, J., 2020. Probabilistic spatiotemporal wind speed forecasting based on a variational Bayesian deep learning model. Appl. Energy 260, 114259.
López, K.L., Gagné, C., Gardner, M.A., 2018. Demand-side management using deep learning for smart charging of electric vehicles. IEEE Trans. Smart Grid 10 (3), 2683–2691.
López, E., Valle, C., Allende, H., Gil, E., 2016. Efficient training over long short-term memory networks for wind speed forecasting. In: Iberoamerican Congress on Pattern Recognition. Springer, Cham, pp. 409–416.
Lydia, M., Suresh Kumar, S., Immanuel Selvakumar, A., Edwin Prem Kumar, G., 2016. Linear and non-linear autoregressive models for short-term wind speed forecasting. Energy Convers. Manage. 112, 115–124.
Lytras, M.D., Chui, K.T., 2019. The recent development of artificial intelligence for smart and sustainable energy systems and applications. Energies 12 (16), 1–7.
Ma, Z., Chen, H., Wang, J., Yang, X., Yan, R., Jia, J., Xu, W., 2020. Application of hybrid model based on double decomposition, error correction and deep learning in short-term wind speed prediction. Energy Convers. Manage. 205, 112345.
Ma, R., Yang, T., Breaz, E., Li, Z., Briois, P., Gao, F., 2018. Data-driven proton exchange membrane fuel cell degradation predication through deep learning method. Appl. Energy 231, 102–115.
Ma, J., Zhang, J., Xiao, L., Chen, K., Wu, J., 2017. Classification of power quality disturbances via deep learning. IETE Technical Review 34 (4), 408–415.
Madan, S., Bollinger, K.E., 1997. Applications of artificial intelligence in power systems. Electr. Power Syst. Res. 41 (2), 117–131.
Mahela, O.P., Shaik, A.G., Gupta, N., 2015. A critical review of detection and classification of power quality events. Renew. Sustain. Energy Rev. 41, 495–505.
Manohar, M., Koley, E., Ghosh, S., 2018. Enhancing the reliability of protection scheme for PV integrated microgrid by discriminating between array faults and symmetrical line faults using sparse auto encoder. IET Renew. Power Gener. 13 (2), 308–317.
Manohar, M., Koley, E., Ghosh, S., Mohanta, D.K., Bansal, R.C., 2020. Spatio-temporal information based protection scheme for PV integrated microgrid under solar irradiance intermittency using deep convolutional neural network. Int. J. Electr. Power Energy Syst. 116, 105576.
Mao, X., Li, Q., Xie, H., Lau, R.Y., Wang, Z., Paul Smolley, S., 2017. Least squares generative adversarial networks. In: Proceedings of the IEEE International.
Marino, D.L., Amarasinghe, K., Manic, M., 2016. Building energy load forecasting using deep neural networks. In: IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society. IEEE, pp. 7046–7051.
Marugán, A.P., Márquez, F.P.G., Perez, J.M.P., Ruiz-Hernández, D., 2018. A survey of artificial neural network in wind energy systems. Appl. Energy 228, 1822–1836.
May, E.C., Ricalde, L.J., Atoche, E.J.R., Bassam, A., Sanchez, E.N., 2018. Forecast and energy management of a microgrid with renewable energy sources using artificial intelligence. In: International Symposium on Intelligent Computing Systems. Springer, Cham, pp. 81–96.
Mbuwir, B., Ruelens, F., Spiessens, F., Deconinck, G., 2017. Reinforcement learning-based battery energy management in a solar microgrid. Energy-Open 2 (4), 36.
Mellit, A., Massi Pavan, A., Ogliari, E., Leva, S., Lughi, V., 2020. Advanced methods for photovoltaic output power forecasting: A review. Appl. Sci. 10 (2), 487.
Memarzadeh, G., Keynia, F., 2020. A new short-term wind speed forecasting method based on fine-tuned LSTM neural network and optimal input sets. Energy Convers. Manage. 213, 112824.
Metaxiotis, K., Kagiannas, A., Askounis, D., Psarras, J., 2003. Artificial intelligence in short term electric load forecasting: a state-of-the-art survey for the researcher. Energy conversion and Management 44 (9), 1525–1534.
Mi, X., Zhao, S., 2020. Wind speed prediction based on singular spectrum analysis and neural network structural learning. Energy Convers. Manage. 216, 112956.
Miikkulainen, R., Liang, J., Meyerson, E., Rawal, A., Fink, D., Francon, O., Raju, B., Shahrzad, H., Navruzyan, A., Duffy, N., Hodjat, B., 2019. Evolving deep neural networks. In: Artificial Intelligence in the Age of Neural Networks and Brain Computing. Academic Press, pp. 293–312.
Mishra, M., 2019. Power quality disturbance detection and classification using signal processing and soft computing techniques: A comprehensive review. International Transactions on Electrical Energy Systems e12008.
Mishra, M., Chandak, S., Rout, P.K., 2019. Taxonomy of islanding detection algorithm for distribution system in microgrid system. Renewable Energy Focus.
Mishra, M., Dash, P.B., Nayak, J., Naik, B., Swain, S.K., 2020. Deep learning and wavelet transform integrated approach for short-term solar PV power prediction. Measurement 108250.
Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., Petersen, S., 2015. Human-level control through deep reinforcement learning. Nature 518 (7540), 529.
Mnih, V., Larochelle, H., Hinton, G.E., 2012. Conditional restricted boltzmann machines for structured output prediction. arXiv preprint arXiv:1202.3748.
Mocanu, D.C., Mocanu, E., Nguyen, P.H., Gibescu, M., Liotta, A., 2016b. Big IoT data mining for real-time energy disaggregation in buildings. In: 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, pp. 003765–003769.

Mocanu, E., Mocanu, D.C., Nguyen, P.H., Liotta, A., Webber, M.E., Gibescu, M., Slootweg, J.G., 2018. On-line building energy optimization using deep reinforcement learning. IEEE Trans. Smart Grid.

Mocanu, E., Nguyen, P.H., Gibescu, M., Kling, W.L., 2016a. Deep learning for estimating building energy consumption. Sustainable Energy, Grids and Networks 6, 91–99.

Mohan, N., Soman, K.P., Vinayakumar, R., 2017. Deep power: Deep learning architectures for power quality disturbances classification. In: 2017 International Conference on Technological Advancements in Power and Energy (TAP Energy). IEEE, pp. 1–6.

Patnaik, B., Mishra, M., Bansal, R.C., Jena, R.K., 2020. AC Microgrid protection–a review: Current and future prospective. Appl. Energy 271, 115210.

Peng, Z., Peng, S., Fu, L., Lu, B., Tang, J., Wang, K., Li, W., 2020. A novel deep learning ensemble model with data denoising for short-term wind speed forecasting. Energy Convers. Manage. 207, 112524.

Potter, C.W., Negnevitsky, M., 2006. Very short-term wind forecasting for tasmanian power generation. IEEE Trans. Power Syst. 21 (2), 965–972.

Qiu, X., Nguyen, T.A., Crow, M.L., 2015. Heterogeneous energy storage optimization for microgrids. IEEE Trans. Smart Grid 7 (3), 1453–1461.

Qiu, X., Ren, Y., Suganthan, P.N., Amaratunga, G.A., 2017. Empirical mode decomposition based ensemble deep learning for load demand time series forecasting. Appl. Soft Comput. 54, 246–255.

Qureshi, A.S., Khan, A., Zameer, A., Usman, A., 2017. Wind power prediction using deep neural network based meta regression and transfer learning. Appl. Soft Comput. 58, 742–755.

Rajiv, K., Tripathi, M.M., 2019. Long short-term memory-convolution neural network based hybrid deep learning approach for power quality events classification. In: Innovations in Electronics and Communication Engineering. Springer, Singapore, pp. 501–510.

Raza, M.Q., Khosravi, A., 2015. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. Renew. Sustain. Energy Rev. 50, 1352–1372.

REN21, 2018. Renewable Energy Policy Network for the 21st Century. Renewable 2018 Global Status Report, REN21.

Ruelens, F., Claessens, B.J., Vandael, S., Iacovella, S., Vingerhoets, P., Belmans, R., 2014. Demand response of a heterogeneous cluster of electric water heaters using batch reinforcement learning. In: 2014 Power Systems Computation Conference. IEEE, pp. 1–7.

Samal, S., Samantaray, S.R., Manikandan, M.S., 2019. A DNN based intelligent protective relaying scheme for microgrids. In: 2019 8th International Conference on Power Systems (ICPS). IEEE, pp. 1–6.

Santamaría-Bonfil, G., Reyes-Ballesteros, A., Gershenson, C., 2016. Wind speed forecasting for wind farms: A method based on support vector regression. Renew. Energy 85, 790–809.

Sergio, A.T., Ludermir, T.B., 2015. Deep learning for wind speed forecasting in northeastern region of Brazil. In: 2015 Brazilian Conference on Intelligent Systems (BRACIS). IEEE, pp. 322–327.

Shekhar, A., Feng, X., Gattozzi, A., Hebner, R., Wardell, D., Strank, S., Rodrigo-Mor, A., Ramírez-Elizondo, L., Bauer, P., 2017. Impact of DC voltage enhancement on partial discharges in medium voltage cables—An empirical study with defects at semicon-dielectric interface. Energies 10 (12), 1968.

Shen, Y., Abubakar, M., Liu, H., Hussain, F., 2019. Power quality disturbance monitoring and classification based on improved PCA and convolution neural network for wind-grid distribution systems. Energies 12 (7), 1280.

Shi, H., Xu, M., Li, R., 2017. Deep learning for household load forecasting—A novel pooling deep RNN. IEEE Trans. Smart Grid 9 (5), 5271–5280.

Soman, S.S., Zareipour, H., Malik, O., Mandal, P., 2010. A review of wind power and wind speed forecasting methods with different time horizons. In: North American Power Symposium 2010. IEEE, pp. 1–8.

Tascikaraoglu, A., Uzunoglu, M., 2014. A review of combined approaches for prediction of short-term wind speed and power. Renew. Sustain. Energy Rev. 34, 243–254.

Taylor, G.W., Hinton, G.E., Roweis, S.T., 2011. Two distributed-state models for generating high-dimensional time series. J. Mach. Learn. Res. 12 (Mar), 1025–1068.

Tong, C., Li, J., Lang, C., Kong, F., Niu, J., Rodrigues, J.J., 2018. An efficient deep model for day-ahead electricity load forecasting with stacked denoising auto-encoders. J. Parallel Distrib. Comput. 117, 267–273.

Tornai, K., Oláh, A., Drenyovszki, R., Kovács, L., Pinté, I., Levendovszky, J., 2017. Recurrent neural network based user classification for smart grids. In: 2017 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT). IEEE, pp. 1–5.

Torres, J.M., Aguilar, R.M., Zuñiga Meneses, K.V., 2018. Deep learning to predict the generation of a wind farm. J. Renew. Sustain. Energy 10 (1), 013305.

Vandael, S., Claessens, B., Ernst, D., Holvoet, T., Deconinck, G., 2015. Reinforcement learning of heuristic EV fleet charging in a day-ahead electricity market. IEEE Trans. Smart Grid 6 (4), 1795–1805.

Vankayala, V.S.S., Rao, N.D., 1993. Artificial neural networks and their applications to power systems—a bibliographical survey. Electr. Power Syst. Res. 28 (1), 67–79.

Venayagamoorthy, G.K., Sharma, R.K., Gautam, P.K., Ahmadi, A., 2016. Dynamic energy management system for a smart microgrid. IEEE Trans. Neural Netw. Learn. Syst. 27 (8), 1643–1656.

Wan, C., Xu, Z., Pinson, P., Dong, Z.Y., Wong, K.P., 2013. Probabilistic forecasting of wind power generation using extreme learning machine. IEEE Trans. Power Syst. 29 (3), 1033–1044.

Wang, S., Chen, H., 2019. A novel deep learning method for the classification of power quality disturbances using deep convolutional neural network. Appl. Energy 235, 1126–1140.

Wang, G., Giannakis, G.B., Chen, J., Sun, J., 2019b. Distribution system state estimation: An overview of recent developments. Front. Inf. Technol. Electron. Eng. 20 (1), 4–17.

Wang, H., Huang, T., Liao, X., Abu-Rub, H., Chen, G., 2016c. Reinforcement learning in energy trading game among smart microgrids. IEEE Trans. Ind. Electron. 63 (8), 5109–5119.

Wang, H., Lei, Z., Zhang, X., Zhou, B., Peng, J., 2019a. A review of deep learning for renewable energy forecasting. Energy Convers. Manage. 198, 111799.

Wang, J., Li, Y., 2018a. Multi-step ahead wind speed prediction based on optimal feature extraction, long short term memory neural network and error correction strategy. Appl. Energy 230, 429–443.

Wang, J., Li, L., Niu, D., et al., 2012. An annual load forecasting model based on support vector regression with differential evolution algorithm. Appl. Energy 94, 65–70.

Wang, H.Z., Li, G.Q., Wang, G.B., Peng, J.C., Jiang, H., Liu, Y.T., 2017a. Deep learning based ensemble approach for probabilistic wind power forecasting. Appl. Energy 188, 56–70.

Wang, Y., Liu, M., Bao, Z., 2016d. Deep learning neural network for power system fault diagnosis. In: 2016 35th Chinese Control Conference (CCC). IEEE, pp. 6678–6683.

Wang, H., Liu, Y., Zhou, B., Li, C., Cao, G., Voropai, N., Barakhtenko, E., 2020. Taxonomy research of artificial intelligence for deterministic solar power forecasting. Energy Convers. Manage. 214, 112909.

Wang, K., Qi, X., Liu, H., Song, J., 2018b. Deep belief network based k-means cluster approach for short-term wind power forecasting. Energy 165, 840–852.

Wang, H., Ruan, J., Wang, G., Zhou, B., Liu, Y., Fu, X., Peng, J., 2018c. Deep learning-based interval state estimation of AC smart grids against sparse cyber attacks. IEEE Trans. Ind. Inf. 14 (11), 4766–4778.

Wang, H.Z., Wang, G.B., Li, G.Q., Peng, J.C., Liu, Y.T., 2016b. Deep belief network based deterministic and probabilistic wind speed forecasting approach. Appl. Energy 182, 80–93.

Wang, H., Yi, H., Peng, J., Wang, G., Liu, Y., Jiang, H., Liu, W., 2017b. Deterministic and probabilistic forecasting of photovoltaic power based on deep convolutional neural network. Energy Convers. Manage. 153, 409–422.

Wang, D., Zhou, Q., Hussain, A., 2016a. Deep and sparse learning in speech and language processing: An overview. In: International Conference on Brain Inspired Cognitive Systems. Springer, Cham, pp. 171–183.

Wani, M.A., Bhat, F.A., Afzal, S., Khan, A.I., 2020. Introduction to deep learning. In: Advances in Deep Learning. Springer, Singapore, pp. 1–11.

Warwick, K., Ekwue, A., Aggarwal, R., Aggarwal, 1997. Artificial Intelligence Techniques in Power Systems (No. 22). Iet.

Wen, Z., O'Neill, D., Maei, H., 2015. Optimal demand response using device-based reinforcement learning. IEEE Trans. Smart Grid 6 (5), 2312–2324.

Wen, L., Zhou, K., Yang, S., 2020. Load demand forecasting of residential buildings using a deep learning model. Electr. Power Syst. Res. 179, 106073.

Wen, L., Zhou, K., Yang, S., Lu, X., 2019. Optimal load dispatch of community microgrid with deep learning based solar power and load forecasting. Energy 171, 1053–1065.

Wu, W., Chen, K., Qiao, Y., Lu, Z., 2016. Probabilistic short-term wind power forecasting based on deep neural networks. In: 2016 International Conference on Probabilistic Methods Applied To Power Systems (PMAPS). IEEE, pp. 1–8.

Wu, H., Liu, Q., Liu, X., 2018. A review on deep learning approaches to image classification and object segmentation. TSP 1 (1), 1–5.

Xie, D., Zhang, L., Bai, L., 2017. Deep learning in visual computing and signal processing. Applied Computational Intelligence and Soft Computing 2017.

Xu, H., Sun, H., Nikovski, D., Kitamura, S., Mori, K., Hashimoto, H., 2019. Deep reinforcement learning for joint bidding and pricing of load serving entity. IEEE Trans. Smart Grid.

Xue, B., Geng, J., 2012. Dynamic transverse correction method of middle and long term energy forecasting based on statistic of forecasting errors. In: 2012 10th International Power & Energy Conference (IPEC). IEEE, pp. 253–256.

Yao, G., Lei, T., Zhong, J., 2019. A review of convolutional-neural-network-based action recognition. Pattern Recognit. Lett. 118, 14–22.

Yoo, Y., Brosch, T., Traboulsee, A., Li, D.K., Tam, R., 2014. Deep learning of image features from unlabeled data for multiple sclerosis lesion segmentation. In: International Workshop on Machine Learning in Medical Imaging. Springer, Cham, pp. 117–124.

You, S.H., Cheng, H.Z., Xie, H., 2006. Mid-and long-term load forecast based on fuzzy linear regression model. Electr. Power Autom. Equip. 26 (3), 51–53.

Yu, R., Gao, J., Yu, M., Lu, W., Xu, T., Zhao, M., Zhang, J., Zhang, R., Zhang, Z., 2019. LSTM-EFG For wind power forecasting based on sequential correlation features. Future Gener. Comput. Syst. 93, 33–42.

Yu, F.R., He, Y., 2019. Reinforcement learning and deep reinforcement learning. In: Deep Reinforcement Learning for Wireless Networks. Springer, Cham, pp. 15–19.

Yue, L., Zhang, Y., Xie, H., et al., 2007. The fuzzy logic clustering neural network approach for middle and long term load forecasting. In: Proc. IEEE Grey Systems and Intelligence Services.

Zafirakis, D., Tzanes, G., Kaldellis, J.K., 2019. Forecasting of wind power generation with the use of artificial neural networks and support vector regression models. Energy Procedia 159, 509–514.

Zang, H., Cheng, L., Ding, T., Cheung, K.W., Liang, Z., Wei, Z., Sun, G., 2018. Hybrid method for short-term photovoltaic power forecasting based on deep convolutional neural network. IET Generation, Transmission & Distribution 12 (20), 4557–4567.

Zang, H., Cheng, L., Ding, T., Cheung, K.W., Wei, Z., Sun, G., 2020. Day-ahead photovoltaic power forecasting approach based on deep convolutional neural networks and meta learning. Int. J. Electr. Power Energy Syst. 118, 105790.

Zazo, J., Zazo, S., Macua, S.V., 2016. Robust worst-case analysis of demand-side management in smart grids. IEEE Trans. Smart Grid 8 (2), 662–673.

Zhang, C.Y., Chen, C.P., Gan, M., Chen, L., 2015. Predictive deep Boltzmann machine for multiperiod wind speed forecasting. IEEE Trans. Sustain. Energy 6 (4), 1416–1425.

Zhang, D., Han, X., Deng, C., 2018. Review on the research and practice of deep learning and reinforcement learning in smart grids. CSEE Journal of Power and Energy Systems 4 (3), 362–370.

Zhang, L., Kekatos, V., Giannakis, G.B., 2016. Scalable electric vehicle charging protocols. IEEE Trans. Power Syst. 32 (2), 1451–1462.

Zhang, Z., Li, S., Xiao, Y., Yang, Y., 2019b. Intelligent simultaneous fault diagnosis for solid oxide fuel cell system based on deep learning. Appl. Energy 233, 930–942.

Zhang, L., Wang, G., Giannakis, G.B., 2019a. Power system state forecasting via deep recurrent neural networks. In: ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, pp. 8092–8096.

Zhao, H., Li, L., 2017. Fault diagnosis of wind turbine bearing based on variational mode decomposition and teager energy operator. IET Renew. Power Gener. 11 (4), 453–460.

Zhao, H., Liu, H., Hu, W., Yan, X., 2018. Anomaly detection and fault analysis of wind turbine components based on deep learning network. Renew. Energy 127, 825–834.

Zhen, Z., Liu, J., Zhang, Z., Wang, F., Chai, H., Yu, Y., Lu, X., Wang, T., Lin, Y., 2020. Deep learning based surface irradiance mapping model for solar PV power forecasting using sky image. IEEE Trans. Ind. Appl..

Zhong, J., Liu, L., Sun, Q., Wang, X., 2018. Prediction of photovoltaic power generation based on general regression and back propagation neural network. Energy Procedia 152, 1224–1229.

Zhou, D.X., 2019. Universality of deep convolutional neural networks. Appl. Comput. Harmon. Anal..

Zhou, J., Liu, H., Xu, Y., Jiang, W., 2018. A hybrid framework for short term multi-step wind speed forecasting based on variational model decomposition and convolutional neural network. Energies 11 (9), 2292.

Zhu, Q., Chen, J., Zhu, L., Duan, X., Liu, Y., 2018. Wind speed prediction with spatio–temporal correlation: A deep learning approach. Energies 11 (4), 705.

Zhu, J., Song, Y., Jiang, D., Song, H., 2017. A new deep-q-learning-based transmission scheduling mechanism for the cognitive internet of things. IEEE Internet Things J. 5 (4), 2375–2385.