Machine Learning Applications for Renewable Energy Systems

Yasir Saleem Afridi, Laiq Hassan, and Kashif Ahmad

1 Introduction

In the modern world, our dependency on technology and modern machines has increased resulting in a significant increase in energy consumption. The World Bank indicators show a strong link between the consumption of energy and the development of a country in different aspects of life, such as economy, infrastructure, health, education, etc. [1]. Fossil fuels, which include crude oil, natural gas, and coal, mainly remained a primary source of energy. However, in the current era, where global warming and climate change are considered to be among the biggest threats to mankind, the focus has been shifted toward renewable sources of energy, such as hydropower, geothermal heat, solar, and wind energy. Compared to fossil fuels, renewable sources of energy bring several advantages. For instance, it is a source of energy that never runs out. More importantly, its zero carbon emission characteristic makes it more environmentally friendly by ensuring cleaner air and water.

While speaking about renewable sources of energy, hydropower has always been the leading source of energy. However, over the past decade, wind power and solar power have also gained a lot of attention [2]. The extensive research in the area has led to the development of technologically advanced and highly complex power generation machines. On one hand, where this development has increased the efficiency and performance of the equipment, it has also generated the need

Y. S. Afridi (✉) · L. Hassan
Department of Computer Systems Engineering, University of Engineering and Technology, Peshawar, Pakistan
e-mail: Yasirsaleem@uetpeshawar.edu.pk

K. Ahmad
Department of Computer Science, Munster Technological University, Cork, Ireland

M. S. Manshahia et al. (eds.), *Advances in Artificial Intelligence for Renewable Energy Systems and Energy Autonomy*, EAI/Springer Innovations in Communication and Computing, https://doi.org/10.1007/978-3-031-26496-2_5

for artificially intelligent energy forecasting, planning, and plant operation and maintenance models/solutions. These models/solutions also make use of state-of-the-art machine learning (ML) algorithms for a diversified set of tasks in the domain.

The current fast-paced advancements in the fields of artificial intelligence (AI) and ML have reduced the need for human intervention in carrying out different complex operations in the power sector to a minimum level. These developments have also enabled us to adopt a proactive approach to the challenges faced while managing such intricate renewable energy systems. The industry experts and analysts foresee a pivotal role being played by AI and ML in the future of renewable energy systems. Investments pertinent to AI in the renewable energy sector are expected to cross USD 7.78 billion by 2024, as per a market intelligence report published by BIS Research [3]. The extensive opportunities for growth offered by AI and ML have forced the major market players to incorporate them into their strategies. The applications of AI and ML in addressing the problems faced by energy companies are plentiful. This chapter provides an overview of AI and ML applications for renewable energy systems. The key applications covered in the chapter include: (i) weather prediction/forecasting using ML algorithms, (ii) forecasting energy supply and demand through AI, (iii) integration of AI with smart grids, and (iv) AI-based condition monitoring and prognostics maintenance systems. An overview of available resources, such as datasets and ML algorithms, for the researchers in the domain is also provided. Moreover, the chapter highlights the key challenges associated with the successful deployment of ML algorithms and potential future research directions in the domain.

The rest of the chapter is organized as follows. Section 2 provides an overview of the existing literature on the topic. Section 3 provides an overview of some key applications of ML for renewable energy systems. Section 5 highlights the key challenges and potential opportunities in the domain. Finally, Sect. 6 concludes the chapter by providing key insights and lessons learned.

2 Related Work

The literature reports outstanding generalization capabilities for ML algorithms in different application domains, such image classification [4], speech recognition [5], and text processing [6]. ML algorithms allow the identification of hidden patterns in a large collection of data and extract meaningful insights. Similar to other application domains, where ML algorithms have been proved very effective, the energy sector also provides a huge amount of data on different aspects of energy systems [2, 7]. To extract meaningful insights from the data, several interesting ML solutions, exploring different aspects of the energy sector, have been proposed in the literature.

Most of the recent efforts in the domain focus on renewable energy systems. For instance, Lai et al. [8] provide a detailed survey of ML models for the prediction tasks in renewable energy systems. The authors discussed different aspects of the domain, such as the performance of existing ML models, pre-processing techniques, and parameter selection approaches adopted in the literature. Gu et al. [9], on the other hand, discuss the use of ML algorithms for renewable energy materials. The authors explored the potential of ML in key renewable energy technologies including catalysis, batteries, solar cells, and crystal discovery. Daniel et al. [10] provide an overview of ML applications in harnessing of renewable energy, such as wind, solar, and thermal energy.

The literature also reports interesting works on certain aspects of ML-based solutions for renewable energy applications. For instance, Salcedo et al. [11] provide an overview of feature selection approaches adopted in ML-based predictive solutions for renewable energy applications. The authors also highlight the challenges, potential, and key aspects of the feature selection process, such as the impact of certain features on the predictive capabilities of ML models in renewable energy applications. Several studies focus on applications of certain types of ML algorithms for renewable energy. For example, Perera et al. [12] explored the potential of reinforcement learning, which represents a sub-category of ML algorithms, in renewable energy applications. In total, seven different applications of renewable energy relying on reinforcement learning algorithms, namely *building energy management system*, *dispatch*, *vehicle energy systems*, *energy devices*, *grids*, *energy markets prediction*, are discussed.

The rich literature on the topic shows the potential of ML in renewable energy applications. However, there are several challenges associated with the successful deployment of ML algorithms in different applications of renewable energy. In this chapter, we highlight such challenges by exploring different aspects of ML applications in renewable energy. We also highlight the potential opportunities, existing resources, and future research directions in the domain.

3 Key Applications

The list of ML applications for renewable energy is very diverse as shown in Fig. 1. In this section, we provide an overview of some of the key applications with a reasonable amount of existing literature, such as weather and energy consumption forecasting, prognostic maintenance, and ML applications in smart grids.

3.1 Forecasting

The process of estimating future events, states, and processes by deploying various conceptual models is known as forecasting. Forecasting is an important aspect of renewable energy systems, specifically solar and wind power, keeping in view their variable energy generation nature. The wind and solar power systems are therefore known as variable renewable energy (VRE) systems because their generation output

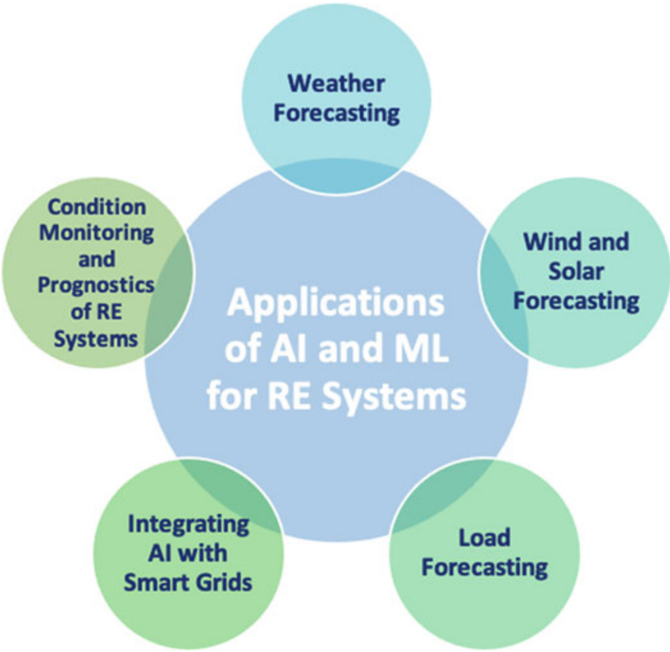

Fig. 1 Some key applications of AI and ML for renewable energy

varies in time, based on the intensity of their sources (i.e., the wind and the Sun). Consequently, an aspect of uncertainty gets associated with them as the power generation by these sources cannot be predicted with perfect accuracy [13]. This is where ML comes into play and is extensively used to carry out forecasts of wind speeds and solar irradiance. In the renewable energy sector, ML algorithms are used for the prediction/forecasting of the future events and states of different elements associated with renewable energy. Irrespective of several forecasting systems being adopted, the model errors continue to exist. However, with the increased use of advanced statistics, ML and AI, the accuracy of these forecasting models has been improved significantly. In the following subsections, we discuss some of the key forecasting applications in the domain.

3.1.1 Weather Forecasting

Weather forecasting plays a vital role in integrating solar and wind power generations into the grid, especially in cases where the penetration levels are high. The most crucial scheduling input for VRE generators that are weather dependent is obtained from weather forecasting data. Therefore, the forecast for power generations is a combination of regional weather forecasts and plant availability.

In the literature, several interesting solutions for weather forecasting have been proposed. These methods can be broadly classified into two categories, namely physical methods and statistical methods [13]. In the physical methods, weather data including temperature, pressure, humidity, surface roughness, and obstacles are fed into a numerical weather predictor (NWP) model. The model in return generates weather conditions using physical and mathematical laws that are specific to terrain and can be converted into energy production. The statistical methods used alongside the NWP models aim to increase the correctness of the results generated by employing historic and real-time generation data. Persistence forecasting, for instance, is used as a benchmark for evaluating the advanced forecasting methods as it is the simplest statistical method based on an assumption that the current generation levels will remain constant in the near future [14]. Advanced forecasting methods make use of AI and big data to carry out the predictions by analyzing live as well as historical weather data. Since advanced forecasting facilitates and improves VRE integration, it is one of the main applications of AI in weather forecasting. With continual advancements in computing power and ML algorithms, these forecasts have become more and more accurate over the past few years.

Moreover, the VRE forecasting approaches could also be categorized into centralized and decentralized methods. Centralized VRE forecasting is the cumulative system-wide forecast of all the VRE generators within a specific balancing area. The centralized forecasts are normally administered by the system operators and are considered to be one of the best approaches for economic dispatch. On the other hand, decentralized VRE forecasting is carried out by individual power generators. It facilitates the system operator in making efficient decisions pertinent to potential transmission congestion by providing plant-level information. The centralized VRE forecasting is more effective as it incorporates a single methodology of forecasting across all distributed power generators [15]. Therefore, it lowers the uncertainties at the system operator level and reduces the financial burden of carrying out individual forecasts on the distributed power generator level. However, relying on a single methodology in the case of centralized forecasting increases the risk of systematic bias. This issue can be addressed by incorporating the ensemble forecasting method, where an aggregate of the results generated by multiple forecast models is taken rather than relying on a single forecasting model [16].

3.1.2 Wind and Solar Power Production Forecasting

In the areas with moderate to high wind power generation, the operators make use of wind energy forecasting to predict the power generation. Likewise, the solar irradiance data are used to forecast solar power forecasting. Given the stochastic nature of wind and solar power generations, ML algorithms have been extensively used for carrying out short-term forecasts of these entities. The current trends in solar and wind power estimation use disaggregation of power generation as well as innovative features and structural information for carrying out short-term forecasts. For instance, based on the video recordings of the sky, convolutional neural networks (CNNs) have been trained to predict near-real-time solar power generation [17]. The solar and wind power output is directly and closely related to the prevailing weather conditions. Recent research, therefore, aims to seamlessly integrate weather forecasting and power generation prediction. Efforts are being made to improve the efficiency of the weather forecasting models to effectively use them as an input for predicting VRE power generation. Numerous physics-based NWP models have been developed to estimate solar irradiance from 0 to 72 h ahead [18]. The output of these multi-timescale NWP models is then used as an input to ML algorithms for carrying out probabilistic power predictions. Similarly, various techniques take the NWP data as features to carry out supervised wind power forecasts. However, the supervised power prediction techniques have a limitation when dealing with distributed energy sources, where the size, location, panel orientation, and hardware data are not always available to the system operators for all the interconnected systems [19]. In such scenarios, satellite and aerial imagery data can be fed into ML models for effectively predicting power outputs [2].

There are several benefits of wind and solar power forecasting. On one hand, it enables the power system operators to maintain lesser reserves. On the other hand, it helps them in handling the supply-side uncertainties. Moreover, efficient predictions can also help in tackling the extreme changes in wind and solar generation that cause a sudden change in the power output. Wind and solar power predictions also enable the grid operators to schedule and dispatch generating plants efficiently. Thereby, the power system operators can make smart and profitable choices on power purchasing by relying more on VRE sources. Furthermore, the wind and solar power prediction also facilitates the power generators by allowing them to carry out plant maintenance during the low production period. Based on the power prediction data, project financers can make better decisions by assessing the plant output data and thus arranging the necessary finance, accordingly.

Nevertheless, the prospects for future research studies to deeply integrate weather and power prediction are quite bright. More studies need to be carried out to develop hybrid physical models where NWP physics-based models are directly incorporated into ML power prediction models.

3.1.3 Load Forecasting

The very fact that electricity cannot be stored in large quantities has given rise to a principle in the power industry known as "the balancing rule." This rule suggests that a consistent balance must always be maintained between the amount of electricity demanded and the amount of electricity supplied. This is important because in both excessive generation and undersupply of electric power, monetary losses are faced by power operators. Hence, to maintain an ideal balancing situation, load forecasting has become an important phenomenon of the current power market. Load forecasting is a technique used by power operators to predict the energy demand to balance it with the anticipated energy supply. Load forecasting also plays a pivotal role in effective decision-making during power system capacity planning to meet the current load requirements and power system expansion to cater the future anticipated load requirements.

Load forecasting can be broadly classified into three groups based on the time horizons of the planning strategies. These categories include: (i) long-term load forecasting (LTLF), (ii) medium-term load forecasting (MTLF), and (iii) short-term load forecasting (STLF). LTLF ranges from 1 year to 20 years and is mainly used for carrying out economic planning of new generation capacity and transmission network. It also facilitates in predicting the future needs for expansion and infrastructure development. MTLF normally predicts the load for a period ranging from 1 week to 1 year. MTLF plays an important role in making decisions pertinent to the scheduling of fuel supplies, carrying out the maintenance activities, financial planning, and tariffs formulation. STLF, on the other hand, ranges from 1 h to 1 week and provides the basis for taking profitable decisions regarding generation units start-up and shutdown. STLFs are used to maintain the demand–supply balance that are important to avoid undersupply and excessive supply of energy. It also provides the information regarding the daily operations and unit commitment to the system operators. STLF is also used to overcome the transmission constraints by providing approximate load flow. STLF also facilitates in economic load dispatching and security assessment.

Load forecasting is a stochastic problem rather than deterministic. Hence, there is no certainty in forecasting. The reason being load forecasting depends upon numerous factors that need to be taken into account while designing a forecasting model. Some of the factors include load density, population growth, historical data, alternative energy sources, and other geo-graphical factors.

Among all the aforementioned load forecasting types, STLF is mostly utilized. One of the potential reasons for more focus on STLF is that it plays a vital role in managing energy transfer schedules based on the estimated load for periods ranging from thirty minutes to an entire day. Therefore, an efficient STLF reduces the expenses incurred by the system operators and enhances the efficiency of the transmission network [20]. In recent years, various techniques have been applied for enhancing the accuracy and efficiency of the load forecasting for VRE systems, including AI and fuzzy logic. Based on their explainability, flexibility to use, and symbolic reasoning, AI has gained more importance and is now being widely used. A technique based on fuzzy logic to carry STLF by incorporating historic weather data has been used by [21]. A detailed review on AI-based load forecasting for smart grids and buildings has been carried out by [22]. ANNs have been used to carry out next day load forecasting [23]. Similarly, an ANN-based STLF for distribution systems [24] and a non-linear autoregressive ANN with exogenous vector inputs to carry out STLF has been developed [25]. Since STLF is based on non-stationary data having forecasting horizon dependencies, LSTM has been used in lieu of its unprecedented ability to handle long-term data dependencies [26]. Atef et al. [27] proposed a deep-stacked LSTM model to forecast the load demand. The results showed that bidirectional LSTM (Bi-LSTM) outperformed the simple LSTM models in terms of forecasting accuracy. Similarly, hybrid ANN models with fuzzy logic have also been developed for accurately predicting the load demand by classifying a large input load dataset. A forecasting model that takes into account the effect of weather and holidays data on the load forecast has been developed using fuzzy logic and ANN [28]. Fuzzy logic has also been used to construct temperature and holiday factor rule bases, an ANN model is then used to predict the hourly load demand. The forecast results showed that the hybrid model produced better results as compared to a standalone ANN model. Also, an enhanced convolutional neural network (CNN) has been proposed to forecast electricity price and load [29]. Numerous AI techniques have been used to carry out load forecasting, and the research show that they have achieved promising results as compared to the conventional techniques. The non-parametric AI-based techniques can clearly overcome the limited capabilities of traditional parametric (statistical) models such as linear regression, stochastic modeling of time series, and general exponential techniques [30]. However, adequate and suitable training data, an appropriate learning algorithm, and an optimized network structure help increasing the overall performance and accuracy of the models, thus reducing the network complexities.

3.2 Integrating AI with Smart Grids

In the past decade, a global paradigm shift from the conventional centralized energy generation to the distributed renewable energy generation has been observed. This has given rise to the need of replacing the traditional transmission and distribution systems with more resilient and smart power distribution and transmission systems. Since, the current grids cannot efficiently cater the fluctuating generation from multiple distributed renewable energy sources and have become obsolete. Therefore, the current grids are now being replaced with the "smart grids." A smart grid is a network that allows two-way flow of data and electricity by effectively integrating digital communication technology with energy distribution, thus, enabling the system operators to optimize the generation, transmission, and distribution of energy on one end and consumers to make cost-effective decisions regarding energy consumption on the other end. Although it is an important factor, but smart grids do not rely on power delivery only, rather the main aspect of a smart grid is a two-way connection of energy and information. Therefore, a smart grid generates an extensive amount of data that is also necessary for its successful operation. Hence, the conventional computational techniques do not have the ability to process such huge amount of data. This is where AI comes into play, capable enough to take into account millions of variables and data points that include but are not limited to weather, location, generation, infrastructure, demand, and assets. AI helps every household and system operator in making proactive decisions regarding the energy generation and supply and the associated energy cost. For instance, if we know ahead of time that it is going to rain for a week, the loss in the solar generation can be catered by proactively upscaling the other generation sources for that specific week. This is what makes AI so appealing for the implementation and management of smart grids. Since the modern power systems are revolutionizing at a fast pace, more and more distributed and diversified smart grid components, such as smart metering systems, digital communication infrastructure, distributed energy sources, and electric vehicles, are getting integrated into the power network along with an underlying communication system. This enables the customers and grid to be directly connected with the help of AI. Thus, homeowners can comprehensively monitor their consumption through smart metering systems and hence take profitable decisions by smartly consuming electricity during low-cost hours. Furthermore, a massive amount of data generated by these smart grid components help automate and enhance the performance of smart grids by supporting vast applications such as forecasting the system state, distributed energy management, fault diagnosis, and grid security against cyberattacks.

3.2.1 Applications of AI in Smart Grids

In smart grids, AI and ML algorithms are used for a diversified set of tasks. In this section, we provide an overview of some of the key applications of AI and ML for smart grids. These applications include:

- **Assessment of Power Grid Stability:** The assessment of the stability of power grids is vital for ensuring the reliability and security of power systems. The power system stability ensures that the system maintains an equilibrium operation state or promptly reaches a new equilibrium state when a small change is induced [31]. The traditional models require extensive computing resources because of their dependence on dynamic power system models. Therefore, data-driven AI-based models are applied to carry out power grid stability analysis because of their efficient performance. The smart grid stability assessment mainly comprises transient stability assessments, small-signal stability, frequency stability, and voltage stability assessment [32]. Transient stability assessment is the ability to determine whether a system will remain synchronized when a huge change in the normal operating state takes place. The small-signal stability assessment, on the other hand, represents the ability of the system to maintain the state of synchronization during small disturbances. The frequency stability assessment is the ability of the system to maintain a steady frequency during the generation and load imbalances, while the voltage stability assessment is the ability of the system to evaluate and maintain voltage stability during a voltage collapse.
- **Faults Diagnosis:** The increased complexity of smart grids has introduced numerous sensitive equipment and components into the system. Protection of such equipment against faults is very important for carrying out smooth operations. Fault diagnosis in smart grids, therefore, provides a defense mechanism for the safety of the sensitive equipment and helps to quickly isolate the faults. With the increased integration of VRE resources in smart grids, effective fault diagnosis has become a great challenge. AI and ML, therefore, play an important

role in carrying out efficient predictive, preventive, and corrective maintenance activities for smart grids.

The literature reports several interesting fault detection techniques. For instance, Fazai et al. [33] proposed an extreme learning machine (ELM) model for fault location detection based on wavelet transform. Similarly, an ensemble framework consisting of five ML algorithms is developed to analyze the power grid frequency disturbances that detected faults with three levels of severity [34]. In [35], a semi-supervised ML model based on KNN and decision tree algorithms are used for fault diagnosis of the transmission and distribution system of microgrids. Extensive research has been carried out in this area that reflects the effectiveness of AI and ML models for carrying out fault diagnosis in smart grids.

- **Security of Smart Grids:** The inherent vulnerability of communication technology and the complexity of smart grids have exposed the communication layers to various security issues. A probable cyberattack on the system can result in operational failures, loss of synchronization, interruption in the power supply, cascading failures, and complete blackouts. Having lethal and vital economic and social consequences, power grids have become a lucrative target for cyberattacks [36]. The most common attacks carried out on smart grids include false data injection attacks (FDIA) and distributed denial of service (DDoS) attacks. In FDIA, the system data are being altered to mislead the power operators, while in DDoS attacks the attackers attempt to make a service unavailable for its intended users. In recent years, various state-of-the-art AI-based approaches have been proposed to ensure the overall security of smart grids. For instance, a neural network model based on stacked denoising autoencoder (SDAE) has been proposed that identifies four different attacks on smart grids [37]. Kosek et al. [38], on the other hand, used an ANN model to identify malicious actions for controlling voltage in the low-voltage distribution grids. Similarly, a semi-supervised ML framework with a domain-adversarial training of known attacks has been used to detect anomalies and patterns for identifying the returning threats at distinct loads and hours [39]. Although sophisticated techniques have been proposed for ensuring the security of the smart grids, however, interdisciplinary research to develop a holistic and methodical solution can further help to tackle the security threats prone to smart grids, effectively.

3.3 Condition Monitoring and Fault Prognostics of Renewable Energy Systems

The continual advancements in the renewable energy sector have led to the development of more complex generation units. Such intricate generation units require more effective and efficient operation and maintenance (O&M) techniques. Furthermore, a delay in diagnosing a fault increases the cost of its rectification. Also, fault can propagate and damage other equipment that can further add to plant shutdown and uncalled-for outages, whereas, with the exponential increase in demand and increased dependency on energy globally, power outages and plant downtimes are highly unfavorable. For ensuring efficient and cost-effective O&M, the current reactive approaches toward fault diagnosis are being replaced with more advanced proactive approaches where the system faults are being predicted. Hence, condition-based monitoring (CBM) and fault prognostics have become the need of the day. To meet these challenges, state-of-the-art AI and ML models have been developed to predict the faults in renewable energy systems, including hydropower, wind power, and solar power projects. In the following subsections, we provide an overview of hydro, wind, and solar projects.

3.3.1 Hydropower Projects

Hydropower plants (HPPs) being one of the first renewable sources of energy are mostly relied upon to cater to the baseload demands of grids. Therefore, plant availability and reduced downtimes are very important while carrying O&M of HPPs. Consequently, the current preventive and corrective maintenance procedures are being replaced with more advanced diagnostic and prognostic maintenance systems. Hence, several efforts are being made for O&M. For instance, a graphical software-based condition monitoring system using wavelet analysis has been developed for a Francis turbine [40]. The research shows that the vast majority of predictive maintenance solutions for HPPs are data-driven. Several data-driven predictive maintenance models for HPPs are being analyzed and classified into three categories: a) physical models, b) stochastic models, and c) ML-based data mining models [41]. Likewise, support vector machine (SVM)-based CBM and fault diagnostic technique for HPPs have been developed [42]. Although SVM outperformed other classification methods, it required higher computational time. Deep learning models have also been used to carry out the CBM and predictive maintenance of HPPs. For instance, a deep neural network-based anomaly detection model in multivariate time-series data has been used [43]. The patterns in the data were captured using long short-term memory (LSTM) because of its unprecedented performance while dealing with time-series data. Similarly, the remaining useful life (RUL) of hydropower turbine bearings has also been determined using the bearing vibrations data acquired from run-to-failure experiments [44]. Although a lot of research have been carried out on the CBM of HPPs; however, variations in the operating conditions of the HPPs make the adaptability (i.e., a model trained on data obtained at one plant could be used for the prediction of data obtained at a plant at a different location with different operating environments) of these models a very complex process.

3.3.2 Wind Power Projects

Among all the renewable energy sources, CBM and fault prognostics for wind turbines (WTs) have always been in the limelight because of the following reasons:

- WTs have got a very high capital cost; therefore, for an effective payback period, plant availability should be maximized.
- WTs normally operate under stressful conditions because of the extreme weather conditions and a constantly variable load; therefore, they have got a higher failure probability.
- Large offshore WTs have got a higher failure downtime because of the access difficulties.

Therefore, an improved CBM and fault prognostics system for WTs can avoid its subassemblies from getting damaged, hence minimizing the plant downtime. Various fascinating and distinct solutions pertinent to the CBM of WTs are reflected in the literature. One of the key components of a WT that transfers the power between the turbine and the generator shaft is the gearbox. Being one of the most critical components, it contributes maximum to the capital cost of the WTs. Consequently, the associated maintenance and repair cost of gearboxes is also very high. Being an important aspect of renewable wind projects, several interesting solutions have been proposed for fault prognostics of wind turbines. For instance, a data-driven framework based on ANN is developed for carrying out the fault prognostics of WT gearbox [45]. The mechanical fault diagnosis of WTs can be further improved by analyzing the vibration signals acquired from the accelerometers along with the power signals. A similar fault diagnostic system based on data mining techniques using multi-sensor data has been proposed [46]. Likewise, models for the estimation of RUL of WTs main bearing have been presented using the likelihood functions [47]. With the advancements in computational resources and the development of new AI algorithms, the search for the development of the most optimal fault diagnostic system for WTs continues.

3.3.3 Solar Power Projects

As compared to wind and hydropower projects, solar power projects are purely electrical and are therefore less susceptible to degradation and faults. Therefore, CBM and fault diagnostics in solar power projects are mostly related to the photovoltaic (PV) modules' health analysis, monitoring the power loss, and the performance monitoring of energy storage systems. For instance, a framework based on ANN is used to carry out the PV health monitoring and analyze the degradation to make effective maintenance decisions accordingly [48]. The performance ratio (PR) of a PV module is the ratio of the actual generation against the rated generation capacity of that module. PR is a key indicator when assessing the reliability of a solar PV system. Various ML techniques have been used to predict the PR of solar power plants to improve the energy reliability [49]. Likewise, performance evaluation of several deep learning techniques including LSTM, ANN, and RNN has been assessed for carrying out the prognosis of solar power projects. The results indicated that LSTM outperformed the other algorithms in terms of accuracy especially while predicting temperature sequences [50]. The research pertinent to

CBM and fault prognostics in solar power projects is still at an inchoate stage and requires further studies to be conducted.

4 Resources (ML Algorithms and Datasets)

4.1 AI/ML Algorithms

In this section, we provide an overview of some of the most commonly used AI/ML algorithms for different applications in renewable energy. For better arrangement, the algorithms are categorized into four categories, namely fuzzy logic, hidden Markov models (HMMs), classical ML algorithms, and neural networks (NNs). In the following subsections, we provide a detailed overview of each of the categories. Moreover, a summary of some of the recent methods from each category is provided in Table 1.

Table 1 Sample works based on each type of ML algorithm discussed in this section

Ref.	ML Model	Application	Description of the method
[54]	FL & NNs	Solar forecasting	It is a hybrid solution combining FL and NNs for solar forecasting. FL is mainly used for pre-processing to correlate key features including cloud cover, wind speed, and temperature
[55]	FL	Fault detection	Relies on FL for comparing electrical parameters against the theoretical parameters to identify faulty PV components
[58]	HMMs	Fault detection	It is a two-step solution where initially PCA is used to extract and select relevant features. An HMM is then trained on the extracted features for the detection and classification of faults
[57]	HMMs	Energy consumption forecasting	Relies on HMMs to deal with the heterogeneous data collected from different sensors for forecasting a day-ahead load
[46]	SVMs	Fault detection	An SVMs classifier, under three different experimental setups, is trained on multi-modal features obtained through different sensors for fault detection in wind turbines
[60]	RF	Energy consumption	Relies on an ensemble of RF classifiers that are trained on features extracted through fast Fourier transform
[66]	CNN & LSTM	Energy state prediction	Relies on a hybrid CNN-LSTM-based framework for energy state prediction from sequences of battery's state of energy and other observable parameters of the mobile edge computing systems
[67]	Power forecasting	CNN & LSTM	Relies on two different hybrid models including a CNN-LSTM and a ConvLSTM trained on uni-variate and multivariate datasets for forecasting power production of a self-consumption PV plant

4.1.1 Fuzzy Logic

Fuzzy logic (FL) represents a subset of AI algorithms that are inspired by the reasoning capabilities of a human. Similar to humans, FL techniques take into account various intermediate possibilities (i.e., degrees of truth) between 0 and 1 [51]. An FL architecture is mainly composed of four components: namely: (i) rules, (ii) fuzzification, (iii) inference engine, and (iv) defuzzification. The first component (i.e., rules) contains a list of rules and conditions (if-then) provided by domain experts. The second component converts inputs (sensor data) into fuzzy sets. The inference engine then determines/decides rules and actions to be performed based on the degree of match between fuzzy input and the rules. Finally, fuzzy sets are converted back into crisp values in the defuzzification process.

FL techniques have been adopted in several types of AI systems in different application domains, such as medicine, autonomous cars/vehicle intelligence, bioinformatics, consumer electronics, and aerospace [52]. FL techniques have also been widely exploited for different applications of AI for renewable energy [2, 53]. Some key applications of FL in the renewable energy sector include prognostics maintenance, site selection for solar power, solar forecasting, and forecasting energy consumption. For instance, Sivaneasan et al. [54] proposed an NNs and FL-based framework for solar forecasting. The FL-based techniques are mainly used to find a correlation of key features, such as cloud cover, temperature, wind speed, and wind direction, with irradiance value. Similarly, Zaki et al. [55] proposed a fault detection framework for solar power systems relying on FL to detect and differentiate eight different types of faults in solar systems. Lau et al. [56], on the other hand, utilize FL for forecasting energy consumption in a manufacturing system. The framework mainly monitors and analyzes the consumption of energy by the manufacturing plant when the functionality/operations of certain production units vary.

4.1.2 Hidden Markov Models (HMMs)

HMMs are state-space models that model the evolution of observable events depending on some non-observable internal factors (hidden states). The observed event is known as a “symbol,” while the non-observable factors are called “states.” We note that HMMs could be used in applications with an observable event “Y” and the non-observable factor/event “X” where the outcome of “Y” is influenced by the outcome of “X” in a known way. In such situations, the goal is to explore the outcome of “X” by observing “Y.” Moreover, the outcome of “Y” at $t = t_0$ must only depend on the outcome/value of “X” at time $t = t_0$, and the outcome of both “Y” and “X” at time $t < t_i$ should not have any impact on the outcome of “Y” at $t = t_0$. This implies that the value of “Y” should not depend on the historical values of “X” at any stage.

The literature reports several variations of HMMs, such as profile-HMMs, maximum entropy Markov models (MEMM), pair-HMMs, and context-sensitive HMMs. HMMs and their variations have been proved very effective in different application domains, such as speech analysis, text recognition, machine translation, and activity recognition. These types of AI algorithms are more effective in applications with sequential and time-series data.

There are several applications of AI in renewable energy that involve analysis of sequential and time-series data, such as predictive maintenance (predicting the remaining useful life of machines), forecasting energy consumption, and load monitoring. The literature already reports several interesting works in this direction. For instance, Bajracharya et al. [57] proposed a HMMs-based energy forecasting framework for predicting a day-ahead load of a data center. The basic motivation behind the proposed solution is to take advantage of HMMs' capabilities of dealing with heterogeneous data in a better way. Kouadri et al. [58], on the other hand, employ HMMs for fault detection in wind energy converter systems. As a first step, the authors use principal component analysis (PCA) to extract and select relevant features. Subsequently, an HMM is trained on the extracted features for the detection and classification of faults.

4.1.3 Conventional ML Algorithms

Conventional ML algorithms, which are also called traditional ML algorithms, represent a subset of ML algorithms that work on features generally extracted by human experts of a domain. These algorithms can be used for several tasks including classification, regression, clustering, and dimensionality reduction. The literature reports a diversified set of traditional ML algorithms, such as support vector machines (SVMs), decision trees, random forest (RF), nearest neighbors, K-means, and Bayes algorithms. Traditional ML algorithms possess several key characteristics that make them a preferable choice for different applications. These characteristics include simplicity in terms of concepts/understanding and implementation. More importantly, these algorithms are interpretable/explainable that bring several advantages to critical and human-centric applications [7]. For instance, interpretable ML models not only result in better failure analysis but also allow an opportunity to further improve the models' performance by tuning them [59].

Most of the initial efforts on intelligent analysis via AI/ML in the energy sector are based on classical ML algorithms [2]. In this regard, the traditional classification algorithms, such as SVMs, RF, and Bayes classifiers, and clustering techniques, such as K-means, self-organizing map (SOM), and Gaussian mixture model (GMM) clustering algorithms, are widely exploited for different tasks in the domain. The key applications of renewable energy where traditional ML algorithms have been shown very effective include prognostic maintenance, energy consumption forecasting, weather forecasting, and other smart grid applications. For instance, Santos et al. [46] employed an SVMs classifier under three different experimental setups for fault detection in wind turbines. The classifier is trained on multi-model features including data from different types of sensors. Similarly, Li et al. [60] rely on RF for forecasting energy consumption.

4.1.4 Artificial Neural Networks (ANNs)

Artificial neural networks (ANNs), which are also called neural networks, represent one of the most commonly used families of AI/ML algorithms. ANNs are mainly inspired by the biological neural system where different algorithms are used to identify hidden patterns in data. Similar to a human brain, ANNs are composed of connected units namely "neurons," which are also known as "nodes" and are based on a mathematical function (i.e., activation function) that collects input data, performs mathematical operations, and produces output according to specific criteria. These neurons are arranged in layers. A typical ANN (i.e., feed-forward NN) generally consists of three types of layers, namely: (i) input, (ii) hidden, and (iii) output layers. ANNs with a single hidden layer are called single-layer perceptrons, while the ANNs with multiple hidden layers are called multi-layer perceptrons.

There are different types of NNs, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and feed-forward NNs, each with a specific set of characteristics. For instance, CNNs are more useful for image analysis. Similarly, RNNs have been proved more effective for sequential and time-series data [4].

In the renewable energy sector, due to the nature of the data, RNNs are most commonly used for different tasks. Among RNNs, long short-term memory (LSTM) and bidirectional LSTM (bi-LSTM) are widely exploited for different tasks in AI applications for renewable energy. Some key applications where LSTM and bi-LSTM have been very effective include prognostic maintenance [2], forecasting energy consumption [61], the impact of climate change on renewable energy resources, risks assessment of renewable resources [62], and forecasting solar power [63]. CNNs are also widely utilized for different applications in the renewable energy sector. CNNs are most used for image-based solutions in the domain. Some key applications of AI for renewable energy where CNNs could be useful include image-based prognostic maintenance of renewable energy systems [2] and power load forecasting [64]. The literature also reports several hybrid solutions combining CNNs and RNNs (LSTM) for different tasks in the domain. These hybrid solutions allow to extract spatio-temporal features that result in better classification performances [65]. For instance, Ku et al. [66] proposed a CNN-LSTM-based solution for predicting the state of energy to avoid battery overcharging and discharging. Similarly, a CNN-LSTM model is proposed by Agga et al. [67] for the prediction of power production of a self-consumption PV plant.

4.2 Datasets

The applicability of ML algorithms in a domain largely depends on the availability of quality data. It is therefore important to provide the readers with a list of publicly available datasets for each application of ML in the renewable energy sector. The literature reports several datasets; however, most of them are not publicly available. Thus, in this section, only publicly available datasets are covered.

4.2.1 Forecasting Energy Supply, Demand, and Weather

Forecasting energy supply and demand is one of the key applications of renewable energy that benefited from ML algorithms. It generally involves processing a huge amount of weather data for meaningful insights, such as weather forecasting, mainly due to the dependency of renewable energy sources on climate changes [68]. Due to this connection, weather data are considered along with other factors for forecasting energy supply and demand. In this section, we discuss some of the publicly available datasets containing energy consumption and weather data for forecasting energy demand and supply via ML algorithms:

- **ENTSOE Dataset [69]**: The dataset provides statistics on electrical consumption, generation, pricing, and weather data for Spain collected from different sources for 4 years. The consumption and energy generation data are obtained from ENTSOE, which is a public portal for transmission service operator (TSO) data. The pricing and weather data, on the other hand, are obtained from the Spanish TSO Red Electric Espana and an open weather API, respectively. One of the key characteristics of the dataset is the hourly consumption data and the corresponding forecasts by the TSO for consumption and pricing, which can be used as a baseline for the underlying ML solutions. Moreover, the objectives of the dataset are multi-fold. For instance, it could be used for: (i) visualization of load and supply data, (ii) analyzing the impact of weather and major cities on the energy supply, demand, and price, and (iii) forecasting hourly and daily energy supply, demand, and price, etc. The dataset is provided in two separate files namely the energy dataset and weather features.
- **Daily Electricity Price and Demand Dataset [70]**: The dataset provides statistics on the daily electricity price and demand in Victoria, Australia. The dataset provides prices and demands for a total of 2016 days from January 1, 2015 to October 6, 2020. The feature set is composed of 14 different features including date/day entries, price, demand, temperature, solar exposure, and rainfall information. For price and demand, it provides values of daily, negative, and positive recommended retail price (RRP). For temperature, both max and min values are provided.
- **Half-hourly Electricity Demand Dataset [71]**: This dataset is also based on the data collected in Victoria, Australia; however, it aims at operational demands. The operational demand represents the demand for energy met by local and semi-scheduled generating units having an aggregated energy higher than 30 MW as well as by energy sources/energy imported to the region. The dataset provides a total of 52,608 data samples/records each containing five different fields of information (i.e., features). These fields include the date, time, electricity

demand in megawatts, temperature, and a binary field indicating public holidays and working days.

- **Building Data Genome 2 (BDG2) Dataset** [72]: It is an open dataset containing data samples (readings) from 3053 different energy meters installed in 1636 different buildings. The data are collected over two years by measuring meter readings on an hourly basis. The dataset provides measurements of electricity consumption, heating, and cooling water, steam, and irrigation meters. The dataset is also used for great energy predictor III (GEPIII) organized by ASHRAE (American Society of Heating and Air-Conditioning Engineers). The dataset also provides weather and other additional information in the form of meta-data. The meta-data file is comprised of 30 different features including key information, such as building ID, site ID, timezone, latitude, longitude, and the number of floors and occupants in the buildings.
- **ISHRAE Weather Dataset [73]**: The dataset covers weather data from 62 different locations in India. The dataset is developed by White Box Technologies by collecting weather data from multiple sources, including the Indian Bureau of Meteorology (IBM), the US National Center for Environmental Data (NCEI), and satellite-derived solar radiation data.
- **Hourly Energy Demand Generation and Weather [74]**: The dataset provides electrical consumption, generation, pricing, and weather data for Spain recorded for 4 years. The data are collected from different sources including ENTSOE (a public portal for Transmission Service Operator (TSO)) and Spanish TSO Red Electric España. The former provides consumption and generation data, while the latter is the source of settlement prices. The weather data, on the other hand, are obtained from the Open Weather API for the 5 largest cities in Spain.
- **Household Electric Power Consumption [75]**: The dataset provides readings of electric power consumption in one household for 4 years. The readings are collected at a sample rate of a one-per minute. It includes readings of different electrical quantities as well as sub-metering values. The feature set includes date, time, global active power, global reactive power, voltage, global intensity, sub-metering (includes readings of kitchen containing a dishwasher, an oven, and a microwave), sub-metering 2 (includes laundry room containing a washing machine, a tumble drier, a refrigerator, and a light), and sub-metering 3 (includes an electric water heater and an air conditioner).

4.2.2 Smart Grids

The literature also provides several datasets to train and evaluate ML algorithms for a diversified set of operations in smart grids. Some of the publicly available datasets in the domain include:

- **Electrical Grid Stability Simulated Dataset [76, 77]**: This is a simulated dataset for local stability analysis of the 4-node star system. The system represents a decentralized smart grid control unit implementing demand response without significant changes to the infrastructure. The dataset provides a total of

10,000 data instances each covering 14 attributes including 11 predictive, 1 non-predictive, and a couple of goal attributes. These attributes include reaction time of energy producers and consumers, power balance (producers and consumers), and price elasticity coefficient (gamma) of energy producers and consumers.
- **SustDataED2 Dataset [78]**: The dataset provides smart meter data, which could be useful to train and evaluate ML algorithms for several applications in smart grids. The dataset provides energy consumption data of individual appliances as well as aggregated consumption of one household in Portugal for 96 days. The data are collected through plug-wise sensors installed at 18 different appliances at 0.5 Hz. Moreover, the data are annotated in a semi-automatic way where first event detection algorithms are used to identify each appliance's events. The events are then manually inspected to verify the labels. The ground truths are provided for both individual appliance consumption and aggregated energy consumption for the house in separate CSV files.

4.2.3 Condition Monitoring and Prognostics Maintenance

Predictive maintenance is one of the key applications of ML for renewable energy systems. In this application, both accelerometer data and endoscopic images can be used. Some publicly available datasets for the application are:

- **Vibration Signal Dataset [79]**: The dataset provides a large collection of data samples (approximately 16,384 instances), generated at a sample rate of 12.8 K samples per second, from six different wind turbines. Although all the samples are generated with the same specification of the wind turbines, the data are organized into six different files each containing data samples generated by a separate turbine. Moreover, the signals are segmented to obtain uniform segments each 1.28 s long. The dataset also provides additional information along with the signal segments. This includes key features, such as the duration for which the data are recorded and the turbine's speed.
- **Wind Generator Dataset [80]**: The dataset is used for the predictive maintenance of wind turbine generators. The dataset provides a diversified set of features covering different aspects of wind turbine generators. In total, each data sample is composed of 101 different features and a single label field with two possible values representing the status of the component either *faulty* or *normal* [80]. The feature set can be roughly divided into environmental conditions (e.g., operational time and wind speed), measurements for wind turbine components (e.g., average rotations per minute), and electrical variables (e.g., voltage, current, and frequency).
- **Gearbox Raw and Elaborated Data [81]:** The dataset is provided in two different forms including a collection of raw gearbox signals and a set of computed features, which is also called elaborated data. Both types of data could be used for predictive maintenance depending on the nature of ML algorithms. The elaborated dataset contains statistical features, such as the standard deviation of accelerations computed at different intervals/frequencies. In the current form,

the elaborated dataset provides data samples at the frequencies of 10, 100, and 1000 consecutive data points. Each data sample is annotated as either a healthy or broken component. We note that the data are generated in a simulated environment through a simulator, namely SpectraQuest by placing four sensors placed at different points. Moreover, the dataset also provides data with different loads ranging from 0 to 90%.

- **Wind Turbine Failure Detection [82]:** The dataset is meant for early-stage failure detection in wind turbines. The dataset provides data on five different components of wind turbines including gearbox, generator, generator bearing, transformer, and hydraulic group. The data are collected through different sensors placed at five different wind turbines for two years at a time interval of 10 min. The resultant dataset is composed of 81 different features including different environmental factors. Moreover, the data are provided in separate training and test set.
- **Grid-connected PV System Faults (GPVS-Faults) Dataset [83]:** This data is also generated in a simulated environment under sixteen different simulation/experimental settings. The data samples generated in each experimental setup are provided in a separate file. Moreover, the dataset provides a deeper annotation hierarchy including: (i) faulty and non-faulty classes (i.e., containing fault-free samples), (ii) types of faults (a total of seven types are covered), and (ii) operational modes, namely limited and maximum power modes. Moreover, the feature set is composed of eleven features including time and various types of current and voltage measurements.

5 Challenges and Open Research Issues

In this section, some of the key open research issues and challenges associated with the successful deployment of ML algorithms for renewable energy applications have been discussed.

- **Availability of Data for Training and Evaluation**: The energy sector is one of the application domains that lack quality annotated data for the training and evaluation of ML algorithms for different tasks. To overcome this issue, the literature reports some efforts for synthesized datasets where data samples are generated in a simulated environment. Though these datasets have been proved effective in training ML models for different tasks, it is very hard to replicate real-life scenarios in a simulated dataset that may affect the performances of the models in real-life applications. Moreover, the generation of these datasets in a simulated environment is also very expensive and requires a lot of effort.
- **Feature Engineering**: In the energy sector, multi-modal data are usually collected through different types of sensors. The selection of the most appropriate and informative features from the heterogeneous data is a challenging and time-consuming job. Moreover, it requires deep knowledge of the domain and a

complete understanding of the data collection process and environmental factors. Though deep-learning-based solutions generally do not involve a feature selection process, it is a critical process for classical ML algorithms. Moreover, the recent shift toward explainable AI solutions has further increased the importance of the feature engineering/selection process [2].

- **Adversarial Attacks**: In the modern world, adversarial attacks, which involve crafting a receivable input sample to misguide or disturb the predictive capabilities of an ML model, are one of the biggest threats to ML-based solutions in critical applications [7, 84]. Renewable energy is one of the critical applications of ML where risks associated with a wrong prediction of an ML model are generally very high. Since most ML models including classical and deep learning models are prone to adversarial attacks, the development of robust ML models for renewable energy applications is the way forward.
- **Integration of Traditional Power Systems in Smart Grids**: The fast-paced development of distributed renewable generation sources and microgrids has resulted in the increased development of smart grids, whereas the traditional power systems still use the old infrastructure for energy distribution. Integration of these traditional power systems in the smart grids has given rise to more uncertainties and complexities for the modern smart grids. This means that smart grids now have to handle an even larger quantity of data, which is still a challenge for them [85]. More research needs to be done to increase the adaptiveness, robustness, and online processing capabilities of the AI algorithms to effectively handle such a large volume of diversified data.
- **Cyberattacks**: As compared to the traditional grids, smart grids opt a two-way communication with multiple integrated devices, which is a lucrative target for cyber attackers. Significant research has been done to develop AI models that can effectively identify the cyber risks; however, smart grids are still prone to a wide variety of attacks [86]. A trade-off is therefore to be made between the performance of AI algorithms and the security of the smart grids.
- **Power Curtailment**: As compared to other sources of energy, VRE such as solar and wind power projects has got relatively low-capacity factor. To meet the demand during peak hours, these projects are over-built in terms of capacity. At times usage or storage of the excess energy is not possible. Therefore, access to energy is being reduced or curtailed. This is not only a monetary wastage but also a wastage of energy. Effective utilization of curtailed energy is still a big challenge for the VRE systems.
- **Load Control**: Currently, the power system operates in a fashion where the load is being adjusted according to the demand of the energy. This limits the dependence on VRE because it is a variable generation. Therefore, the need for curtailment and reduced generation is also there. This challenge can be addressed if the power systems are designed in a way that instead of matching the load, energy shall be utilized only when there is a VRE generation in the grid. A paradigm shift from load control to demand control is therefore needed. Hence, controllable and responsive loads being one of the most underutilized reliability resources can balance the demand and supply over all time frames ranging from seconds to seasons.

6 Conclusions

This chapter discusses the key applications of ML and AI for renewable energy by providing a detailed overview of challenges, available resources in the form of datasets and ML algorithms, and potential future research directions. The chapter discusses how AI and ML algorithms can help in forecasting future events, states, and processes associated with renewable energy. Moreover, an overview of some of the key applications of AI and ML in smart grids and prognostic maintenance is also provided. The literature shows that AI and ML can play a vital role in further enhancing the productivity and management of renewable energy resources. Despite being widely explored over the last few years, several aspects need to be considered to fully explore the potential of AI and ML in the renewable energy sector.

References

1. Lloyd, P. J. (2017). The role of energy in development. *Journal of Energy in Southern Africa, 28*(1), 54–62.
2. Afridi, Y. S., Ahmad, K., & Hassan, L. (2021). Artificial intelligence based prognostic maintenance of renewable energy systems: A review of techniques, challenges, and future research directions. *International Journal of Energy Research*.
3. Global artificial intelligence (AI) in energy market. https://bisresearch.com/industry-report/artificial-intelligence-energy-market.html
4. Ahmad, K., & Conci, N. (2019). How deep features have improved event recognition in multimedia: A survey. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 15*(2), 1–27.
5. Bianco, M. J., Gerstoft, P., Traer, J., Ozanich, E., Roch, M. A., Gannot, S., & Deledalle, C.-A. (2019). Machine learning in acoustics: Theory and applications. *The Journal of the Acoustical Society of America, 146*(5), 3590–3628.
6. Kadhim, A. I. (2019). Survey on supervised machine learning techniques for automatic text classification. *Artificial Intelligence Review, 52*(1), 273–292.
7. Ahmad, K., Maabreh, M., Ghaly, M., Khan, K., Qadir, J., & Al-Fuqaha, A. (2022). Developing future human-centered smart cities: Critical analysis of smart city security, data management, and ethical challenges. *Computer Science Review, 43*, 100452.
8. Lai, J.-P., Chang, Y.-M., Chen, C.-H., & Pai, P.-F. (2020). A survey of machine learning models in renewable energy predictions. *Applied Sciences, 10*(17), 5975.
9. Gu, G. H., Noh, J., Kim, I., & Jung, Y. (2019). Machine learning for renewable energy materials. *Journal of Materials Chemistry A, 7*(29), 17096–17117.
10. Daniel, C., Shukla, A. K., & Sharma, M. (2021). Applications of machine learning in harnessing of renewable energy. In *Advances in clean energy technologies* (pp. 177–187). Springer.
11. Salcedo-Sanz, S., Cornejo-Bueno, L., Prieto, L., Paredes, D., & Garcia-Herrera, R. (2018). Feature selection in machine learning prediction systems for renewable energy applications. *Renewable and Sustainable Energy Reviews, 90*, 728–741.
12. Perera, A., & Kamalaruban, P. (2021). Applications of reinforcement learning in energy systems. *Renewable and Sustainable Energy Reviews, 137*, 110618.
13. Meenal, R., Binu, D., Ramya, K., Michael, P. A., Vinoth Kumar, K., Rajasekaran, E., & Sangeetha, B. (2022). Weather forecasting for renewable energy system: A review. *Archives of Computational Methods in Engineering, 29*(5), 2875–2891.

14. Monteiro, C., Keko, H., Bessa, R., Miranda, V., Botterud, A., Wang, J., & Conzelmann, G. (2009). A quick guide to wind power forecasting: state-of-the-art 2009. Technical report, Argonne National Lab. (ANL), Argonne, IL (United States).
15. Katz, J. (2015). Grid integration studies: Data requirements, greening the grid. Technical report, National Renewable Energy Lab. (NREL), Golden, CO (United States).
16. Chernyakhovskiy, I., et al. (2016). Forecasting wind and solar generation: Improving system operations, greening the grid. Technical report, National Renewable Energy Lab. (NREL), Golden, CO (United States).
17. Sun, Y., Szűcs, G., & Brandt, A. R. (2018). Solar PV output prediction from video streams using convolutional neural networks. *Energy & Environmental Science, 11*(7), 1811–1818.
18. Yang, D. (2019). On post-processing day-ahead NWP forecasts using Kalman filtering. *Solar Energy, 182*, 179–181.
19. Foley, A. M., Leahy, P. G., Marvuglia, A., & McKeogh, E. J. (2012). Current methods and advances in forecasting of wind power generation. *Renewable Energy, 37*(1), 1–8.
20. Sadaei, H. J., e Silva, P. C. d. L., Guimarães, F. G., & Lee, M. H. (2019). Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series. *Energy, 175*, 365–377.
21. Ranaweera, D., Hubele, N., & Karady, G. (1996). Fuzzy logic for short term load forecasting. *International Journal of Electrical Power & Energy Systems, 18*(4), 215–222.
22. Raza, M. Q., & Khosravi, A. (2015). A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. *Renewable and Sustainable Energy Reviews, 50*, 1352–1372.
23. Velasco, L. C. P., Villezas, C. R., Palahang, P. N. C., & Dagaang, J. A. A. (2015). Next day electric load forecasting using artificial neural networks. In *2015 International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)* (pp. 1–6). IEEE.
24. Hernández, L., Baladrón, C., Aguiar, J. M., Calavia, L., Carro, B., Sánchez-Esguevillas, A., Pérez, F., Fernández, Á., & Lloret, J. (2014). Artificial neural network for short-term load forecasting in distribution systems. *Energies, 7*(3), 1576–1598.
25. Buitrago, J., & Asfour, S. (2017). Short-term forecasting of electric loads using nonlinear autoregressive artificial neural networks with exogenous vector inputs. *Energies, 10*(1), 40.
26. Rafi, S. H., et al. (2020). Highly efficient short term load forecasting scheme using long short term memory network. In *2020 8th International Electrical Engineering Congress (iEECON)* (pp. 1–4). IEEE.
27. Atef, S., & Eltawil, A. B. (2020). Assessment of stacked unidirectional and bidirectional long short-term memory networks for electricity load forecasting. *Electric Power Systems Research, 187*, 106489.
28. Yang, K., & Zhao, L. (2009). Application of Mamdani fuzzy system amendment on load forecasting model. In *2009 Symposium on Photonics and Optoelectronics* (pp. 1–4). IEEE.
29. Zahid, M., Ahmed, F., Javaid, N., Abbasi, R. A., Zainab Kazmi, H. S., Javaid, A., Bilal, M., Akbar, M., & Ilahi, M. (2019). Electricity price and load forecasting using enhanced convolutional neural network and enhanced support vector regression in smart grids. *Electronics, 8*(2), 122.
30. Baliyan, A., Gaurav, K., & Mishra, S. K. (2015). A review of short term load forecasting using artificial neural network models. *Procedia Computer Science, 48*, 121–125.
31. Baltas, N. G., Mazidi, P., Ma, J., de Asis Fernandez, F., & Rodriguez, P. (2018). A comparative analysis of decision trees, support vector machines and artificial neural networks for on-line transient stability assessment. In *2018 International Conference on Smart Energy Systems and Technologies (SEST)* (pp. 1–6). IEEE.
32. You, S., Zhao, Y., Mandich, M., Cui, Y., Li, H., Xiao, H., Fabus, S., Su, Y., Liu, Y., Yuan, H., et al. (2020). A review on artificial intelligence for grid stability assessment. In *2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)* (pp. 1–6). IEEE.

33. Fazai, R., Abodayeh, K., Mansouri, M., Trabelsi, M., Nounou, H., Nounou, M., & Georghiou, G. E. (2019). Machine learning-based statistical testing hypothesis for fault detection in photovoltaic systems. *Solar Energy, 190*, 405–413.
34. Niu, H., Omitaomu, O. A., & Cao, Q. C. (2020). Machine committee framework for power grid disturbances analysis using synchrophasors data. *Smart Cities, 4*(1), 1–16.
35. Abdelgayed, T. S., Morsi, W. G., & Sidhu, T. S. (2017). Fault detection and classification based on co-training of semisupervised machine learning. *IEEE Transactions on Industrial Electronics, 65*(2), 1595–1605.
36. Tufail, S., Parvez, I., Batool, S., & Sarwat, A. (2021). A survey on cybersecurity challenges, detection, and mitigation techniques for the smart grid. *Energies, 14*(18), 5894.
37. Zhou, L., Ouyang, X., Ying, H., Han, L., Cheng, Y., & Zhang, T. (2018). Cyber-attack classification in smart grid via deep neural network. In *Proceedings of the 2nd International Conference on Computer Science and Application Engineering* (pp. 1–5).
38. Kosek, A. M. (2016). Contextual anomaly detection for cyber-physical security in smart grids based on an artificial neural network model. In *2016 Joint Workshop on Cyber-Physical Security and Resilience in Smart Grids (CPSR-SG)* (pp. 1–6). IEEE.
39. Zhang, Y., & Yan, J. (2020). Semi-supervised domain-adversarial training for intrusion detection against false data injection in the smart grid. In *2020 International Joint Conference on Neural Networks (IJCNN)* (pp. 1–7). IEEE.
40. Liu, Z., Zou, S., & Zhou, L. (2012). Condition monitoring system for hydro turbines based on LabVIEW. In *2012 Asia-Pacific Power and Energy Engineering Conference* (pp. 1–4). IEEE.
41. Wang, S., Wang, K., & Li, Z. (2016). A review on data-driven predictive maintenance approach for hydro turbines/generators. In *6th International Workshop of Advanced Manufacturing and Automation* (pp. 30–35). Atlantis Press.
42. Selak, L., Butala, P., & Sluga, A. (2014). Condition monitoring and fault diagnostics for hydropower plants. *Computers in Industry, 65*(6), 924–936.
43. Zhang, C., Song, D., Chen, Y., Feng, X., Lumezanu, C., Cheng, W., Ni, J., Zong, B., Chen, H., & Chawla, N. V. (2019). A deep neural network for unsupervised anomaly detection and diagnosis in multivariate time series data. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 33, pp. 1409–1416).
44. Wang, D., Tsui, K.-L., & Miao, Q. (2017). Prognostics and health management: A review of vibration based bearing and gear health indicators. *IEEE Access, 6*, 665–676.
45. Verma, A., Zappalá, D., Sheng, S., & Watson, S. J. (2022). Wind turbine gearbox fault prognosis using high-frequency SCADA data. In *Journal of Physics: Conference Series* (Vol. 2265, p. 032067). IOP Publishing.
46. Santos, P., Villa, L. F., Reñones, A., Bustillo, A., & Maudes, J. (2015). An SVM-based solution for fault detection in wind turbines. *Sensors, 15*(3), 5627–5648.
47. Herp, J., Pedersen, N. L., & Nadimi, E. S. (2019). Assessment of early stopping through statistical health prognostic models for empirical RUL estimation in wind turbine main bearing failure monitoring. *Energies, 13*(1), 83.
48. Riley, D., & Johnson, J. (2012). Photovoltaic prognostics and heath management using learning algorithms. In *2012 38th IEEE Photovoltaic Specialists Conference* (pp. 001535–001539). IEEE.
49. Bandong, S., Leksono, E., Purwarianti, A., & Joelianto, E. (2019). Performance ratio estimation and prediction of solar power plants using machine learning to improve energy reliability. In *2019 6th International Conference on Instrumentation, Control, and Automation (ICA)* (pp. 36–41). IEEE.
50. Correa-Jullian, C., Cardemil, J. M., Droguett, E. L., & Behzad, M. (2020). Assessment of deep learning techniques for prognosis of solar thermal systems. *Renewable Energy, 145*, 2178–2191.
51. Dernoncourt, F., & Sander, E. (2011). Fuzzy logic: between human reasoning and artificial intelligence. *Project Report Paris*.
52. Abduljabbar, R., Dia, H., Liyanage, S., & Bagloee, S. A. (2019). Applications of artificial intelligence in transport: An overview. *Sustainability, 11*(1), 189.

53. Suganthi, L., Iniyan, S., & Samuel, A. A. (2015). Applications of fuzzy logic in renewable energy systems—a review. *Renewable and Sustainable Energy Reviews, 48*, 585–607.
54. Sivaneasan, B., Yu, C., & Goh, K. (2017). Solar forecasting using ANN with fuzzy logic pre-processing. *Energy Procedia, 143*, 727–732.
55. Zaki, S. A., Zhu, H., & Yao, J. (2019). Fault detection and diagnosis of photovoltaic system using fuzzy logic control. In *E3S Web of Conferences* (Vol. 107, p. 02001). EDP Sciences.
56. Lau, H. C., Cheng, E., Lee, C. K., & Ho, G. T. (2008). A fuzzy logic approach to forecast energy consumption change in a manufacturing system. *Expert Systems with Applications, 34*(3), 1813–1824.
57. Bajracharya, A., Khan, M. R. A., Michael, S., & Tonkoski, R. (2018). Forecasting data center load using hidden Markov model. In *2018 North American Power Symposium (NAPS)* (pp. 1–5). IEEE.
58. Kouadri, A., Hajji, M., Harkat, M.-F., Abodayeh, K., Mansouri, M., Nounou, H., & Nounou, M. (2020). Hidden Markov model based principal component analysis for intelligent fault diagnosis of wind energy converter systems. *Renewable Energy, 150*, 598–606.
59. Khan, T., Ahmad, K., Khan, J., Khan, I., & Ahmad, N. (2022). An explainable regression framework for predicting remaining useful life of machines. arXiv preprint arXiv:2204.13574.
60. Li, C., Tao, Y., Ao, W., Yang, S., & Bai, Y. (2018). Improving forecasting accuracy of daily enterprise electricity consumption using a random forest based on ensemble empirical mode decomposition. *Energy, 165*, 1220–1227.
61. Yan, K., Li, W., Ji, Z., Qi, M., & Du, Y. (2019). A hybrid LSTM neural network for energy consumption forecasting of individual households. *IEEE Access, 7*, 157633–157642.
62. Liu, B., Chen, J., Wang, H., & Wang, Q. (2020). Renewable energy and material supply risks: A predictive analysis based on an LSTM model. *Frontiers in Energy Research, 8*, 163.
63. 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)* (pp. 002858–002865). IEEE.
64. Aurangzeb, K., Alhussein, M., Javaid, K., & Haider, S. I. (2021). A pyramid-CNN based deep learning model for power load forecasting of similar-profile energy customers based on clustering. *IEEE Access, 9*, 14992–15003.
65. Kumari, P., & Toshniwal, D. (2021). Long short term memory–convolutional neural network based deep hybrid approach for solar irradiance forecasting. *Applied Energy, 295*, 117061.
66. Ku, Y.-J., Sapra, S., Baidya, S., & Dey, S. (2020). State of energy prediction in renewable energy-driven mobile edge computing using CNN-LSTM networks. In *2020 IEEE Green Energy and Smart Systems Conference (IGESSC)* (pp. 1–7). IEEE.
67. Agga, A., Abbou, A., Labbadi, M., & El Houm, Y. (2021). Short-term self consumption PV plant power production forecasts based on hybrid CNN-LSTM, ConvLSTM models. *Renewable Energy, 177*, 101–112.
68. Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., et al. (2022). Tackling climate change with machine learning. *ACM Computing Surveys (CSUR), 55*(2), 1–96.
69. Hourly energy demand generation and weather. https://www.kaggle.com/datasets/nicholasjhana/energy-consumption-generation-prices-and-weather
70. Daily electricity price and demand data. https://www.kaggle.com/datasets/aramacus/electricity-demand-in-victoria-australia
71. Half-hourly electricity demand for Victoria, Australia. https://tsibbledata.tidyverts.org/reference/vic_elec.html
72. Miller, C., Kathirgamanathan, A., Picchetti, B., Arjunan, P., Park, J. Y., Nagy, Z., Raftery, P., Hobson, B. W., Shi, Z., & Meggers, F. (2020). The Building Data Genome Project 2, energy meter data from the ASHRAE great energy predictor III competition. *Scientific Data, 7*, 368.
73. ISHAREA dataset. http://weather.whiteboxtechnologies.com/ISHRAE
74. Hourly energy demand generation and weather dataset. https://www.kaggle.com/datasets/nicholasjhana/energy-consumption-generation-prices-and-weather

75. Household electric power consumption dataset. https://www.kaggle.com/datasets/uciml/electric-power-consumption-data-set
76. Smard grid stability dataset. https://tinyurl.com/2t2ddhm8
77. Arzamasov, V., Bohm, K., & Jochem, P. (2018). Towards concise models of grid stability. In *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)* (pp. 1–6). IEEE.
78. Pereira, L., Costa, D., & Ribeiro, M. (2022). A residential labeled dataset for smart meter data analytics. *Scientific Data, 9*(1), 1–11.
79. Martin del Campo Barraza, S., Sandin, F., & Strömbergsson, D. (2018). Dataset concerning the vibration signals from wind turbines in northern Sweden.
80. Mammadov, E. E. O. (2019). Predictive maintenance of wind generators based on AI techniques. Master's thesis, University of Waterloo.
81. Gearbox fault diagnosis: elaborated datasets. https://tinyurl.com/mbu5cv8w
82. The EDP dataset. https://tinyurl.com/2nva3ta5
83. GPVS-faults: Experimental data for fault scenarios in grid-connected PV systems under MPPT and IPPT modes. https://tinyurl.com/nw2aha8
84. Al-Maliki, S., El Bouanani, F., Ahmad, K., Abdallah, M., Hoang, D., Niyato, D., & Al-Fuqaha, A. (2021). Opportunistic use of crowdsourced workers for online relabeling of potential adversarial examples.
85. Kazmi, S. A. A., Hasan, S. F., & Shin, D.-R. (2015). Analyzing the integration of distributed generation into smart grids. In *2015 IEEE 10th Conference on Industrial Electronics and Applications (ICIEA)* (pp. 762–766). IEEE.
86. Gunduz, M. Z., & Das, R. (2018). Analysis of cyber-attacks on smart grid applications. In *2018 International Conference on Artificial Intelligence and Data Processing (IDAP)* (pp. 1–5). IEEE.