Renewable and Sustainable Energy Reviews 160 (2022) 112128

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

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

Data-driven probabilistic machine learning in sustainable smart energy/smart energy systems: Key developments, challenges, and future research opportunities in the context of smart grid paradigm

Tanveer Ahmad [a,*], Rafal Madonski [a], Dongdong Zhang [b], Chao Huang [c,d], Asad Mujeeb [e]

[a] Energy and Electricity Research Center, International Energy College, Jinan University, Zhuhai, Guangdong Province, 519070, China
[b] School of Electrical Engineering, Guangxi University, Nanning, 530004, China
[c] School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing, 10083, China
[d] Shunde Graduate School, University of Science and Technology Beijing, Guangdong, 528399, China
[e] Department of Electrical Engineering, Tsinghua University, Beijing, 100084, PR China

A R T I C L E I N F O

Keywords:
Data-driven probabilistic machine learning
Energy distribution
Discovery and design of energy materials
Big data analytics and smart grid
Strategic energy planning and smart manufacturing
Energy demand-side response

A B S T R A C T

The current trend indicates that energy demand and supply will eventually be controlled by autonomous software that optimizes decision-making and energy distribution operations. New state-of-the-art machine learning (ML) technologies are integral in optimizing decision-making in energy distribution networks and systems. This study was conducted on data-driven probabilistic ML techniques and their real-time applications to smart energy systems and networks to highlight the urgency of this area of research. This study focused on two key areas: i) the use of ML in core energy technologies and ii) the use cases of ML for energy distribution utilities. The core energy technologies include the use of ML in advanced energy materials, energy systems and storage devices, energy efficiency, smart energy material manufacturing in the smart grid paradigm, strategic energy planning, integration of renewable energy, and big data analytics in the smart grid environment. The investigated ML area in energy distribution systems includes energy consumption and price forecasting, the merit order of energy price forecasting, and the consumer lifetime value. Cybersecurity topics for power delivery and utilization, grid edge systems and distributed energy resources, power transmission, and distribution systems are also briefly studied. The primary goal of this work was to identify common issues useful in future studies on ML for smooth energy distribution operations. This study was concluded with many energy perspectives on significant opportunities and challenges. It is noted that if the smart ML automation is used in its targeting energy systems, the utility sector and energy industry could potentially save from $237 billion up to $813 billion.

1. Introduction

Today, while countries seek to restructure their energy strategies and make cleaner energy more dependent, one major challenge remains [1]. Both wind and solar power are, by definition, intermittent nature of sources of electricity [2]. The power output of a solar panel or wind turbine is never constant; it is determined by external variables such as cloud cover intensity, solar radiation, and wind speed, all of which are uncontrollable [3]. When wind and solar farms generate less electricity, grid operators must switch to traditional power plants to balance energy supply and demand. Alternatively, after 90% of their electricity needs for the day have been met on windy and sunny days, operators must decrease gas-fired and coal power plant production to prevent a power overload allowing the whole grid to suffocate or be damaged. Grid operators have to compensate energy providers for making changes to their supplies (i.e., power system infrastructure), saving German consumers, for example, about $553 million a year [4].

Besides, electric utilities generate excessive carbon dioxide emissions when their excess electricity is dissipated [5]. All this contributes to the complexities of accurately forecasting the industry's health status of energy distribution infrastructure. In general, keeping the demand and supply of energy balanced will turn into a constant operational and technological struggle. This leads us to the potential that ML has and to the major impact it could have on the energy spectrum as a whole [6]. Even though ML is in its early stages of deployment, it could fundamentally transform how we communicate with resources [7]. Its effect could be far-reaching in renewable energy distribution and forecasts and

* Corresponding author.
E-mail addresses: tanveer.ahmad.pk11@gmail.com, tanveerahma@jnu.edu.cn (T. Ahmad).

https://doi.org/10.1016/j.rser.2022.112128
Received 23 March 2021; Received in revised form 5 December 2021; Accepted 10 January 2022
Available online 5 March 2022

Abbreviations

ML	Machine learning
TWh	Terawatt-hours
REmap	Renewable energy map
PV	Photovoltaics
ROI	Return on investment
DOE	Department of Energy
IBM	International Business Machines
ABB	ASEA Brown Boveri
BP	British Petroleum
AI	Artificial intelligence
SVM	Support vector machine
CNN	Convolution neural networks
DBNN	Deep belief neural networks
DBM	Deep Boltzmann machine
RNN	Recurrent neural networks
DRL	Deep reinforcement learning
LoRa	Short for long-range
PCC	Point of common coupling
HEM	Home energy management
EV	Electric vehicles
ESS	Energy storage system
RES	Renewable energy source
US	The United States of America
AEPC	Active building energy performance contracting
CO2	Carbon dioxide
USD	The United States dollar
IoT	Internet of Things
SCADA	Supervisory control and data acquisition control system
RTUs	Remote terminal unit
GIS	Geographic information system
CLV/LTV	Consumer lifetime value
TEASER	Energy analysis and simulation for efficient retrofit
RL	Reinforcement learning
EDLC	Electric double-layer capacitors
SMES	Superconducting magnetic energy storage
V2G	Vehicle-to-grid
TESS	Thermal energy storage system

the implementation of smart grids.

The aim and objective of this study: Right now, ML is considered a "hot topic" in a variety of research fields, and it is currently the most rapidly developing sector in high-tech. Data processing and interpretation are becoming increasingly important, even for engineers, due to energy digitalization. Smart grids are a term used to describe various new data-based services in renewable energy supply, marketing, storage, and usage. This study covers recent advances and fundamental ML techniques in energy distribution. ML assists in the fast and efficient processing of this data in energy distribution (e.g., the energy distribution is the last stage of energy delivery, it carries out the energy from transmission systems, conventional grid or smart grid infrastructure, individual consumers, transmission and distribution infrastructure, energy devices and materials, big data analytics, etc.). To start, we identify the numerous potential challenges that ML tries to solve, review recent advances in the field, and analyze ML's effect on the energy sector. The introduction of different ML model classes to solve such complex problems is briefly examined. Eight various aspects of ML in core energy technologies are discussed. Five use cases of ML in energy distribution are briefly covered. The used software for ML models in energy distribution, challenges, ML opportunities towards a smart and sustainable future, and recent progress on discovery and properties of ML models aim to strengthen the core theme of this study.

The Framework and Structure of this Study: The description of the performed study has been arranged into eight sections. The paper opens with a general introduction of the study, including four subsections. The process of energy conversion (supporting shift towards 100% renewable energy) is discussed in the first subsection. The main idea and concept of ML, historical overview of ML methods, list of ML models and their usage in the smart energy systems, as well as the role of ML in the distribution of energy are described in Sects. 1.2 to 1.5, respectively. Then, Sect. 2 follows the core concept of energy technologies using ML, which discusses the use of ML in various areas. The use of ML for energy distribution utilities is elaborated in Sect. 3, with its five subsections dedicated to accurate energy price forecasting, keeping the merit order of accurate energy prices, predicting consumer life values, probability assessment of winning consumers, and making good offers to the energy consumers. Sect. 4 highlights various uses of ML tools for energy distribution systems and introduces different software used by organizations that deal with ML techniques. Sect. 5 describes the analysis of the existing challenges of ML in the context of energy distribution systems. We then go on to identify ML opportunities towards a smart and sustainable future in Sect. 6. Recent progress on properties and discovery of ML is described in Sect. 7. Finally, Sect. 8 concluded the conducted study.

Fig. 1 shows the results of an analysis of search queries on ML in the subject area, article type, and area of research in engineering and energy distribution. The analysis is divided into three major parts: including i) document by subject area on ML models (e.g., published research articles in all areas of research but with the use of ML models), ii) document by subject area on ML in engineering (e.g., published research articles in the engineering field); iii) and document by subject area on ML in engineering and energy (e.g., published research articles in the engineering field, particularly in energy). Overall, the United States is leading in publishing research articles in all fields. China is number one in the World for publishing research articles in the engineering and energy fields. Fig. 1 shows the data clusters result from Scopus.

1.1. Energy conversation: moving 100% towards clean and renewable energy

Globally, a total of 376 TW-hours (TWh) of renewable energy was produced in 2018, a rise of 6.1% relative to the previous year (2017) [8]. Wind and solar production rose by 11% and 28%, respectively, in 2018 [8]. In 2018, Asia accounted for much of the increase in renewable energy production, increasing generation by a total of 219 TWh. Asia's volume of world renewable generation also rose, increasing to 40% of overall renewable generation. North America and Europe together hold 40% of the shares, led by Eurasia (5%) and South America (12%) [8].

The United States' clean energy initiative is becoming more assertive: more groups target a rise in renewable energy use and carbon emissions decrease [9]. Many states have recently declared a target to increase their current renewable portfolio requirements to incorporate a 100% renewable energy standard (e.g., New Mexico, Washington, Nevada, and Colorado) [10]. A rising number of power companies are now making enormous commitments to reduce greenhouse emissions and increase sustainable energy usage [11]. At least ten publicly listed utility companies have declared 100% decarbonization targets, and a large portion of these utilities have established goals to minimize carbon

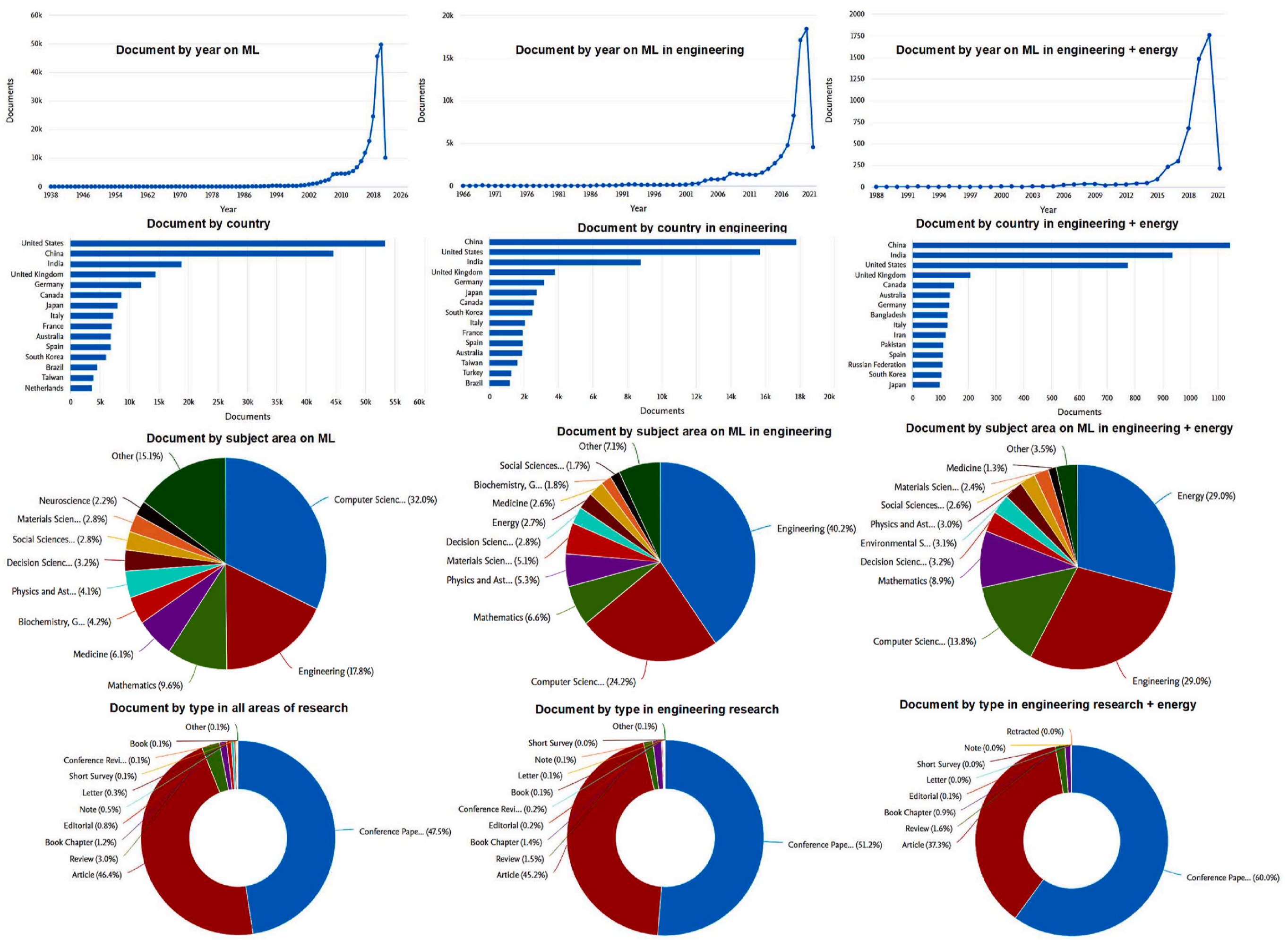


Fig. 1. ML search results in engineering and energy distribution were examined in terms of subject, article type, and research area.

emissions by much more than 80% below 2005 levels by 2050 [12].

Fig. 2 demonstrates[1] the proportion of the various renewable energy sources in overall final energy use worldwide in 2010 as well as for 2030 with the additional increase arising from the renewable energy map (REmap) opportunities [13]. The primary source of green energies would be bioenergy, which can produce both heat and power and vehicle fuels. By 20% of renewable energy usage in REmap 2030, many natural resources of liquid, solid, and gaseous biomass constitute 61% of the total [13]. However, as previously reported, most of the transition would be the move from conventional to modern technologies and fuels [14]. By 2030, REmap strategies would boost both the total volume and share of wind power usage, while wind power deployment would surpass hydropower. Solar-PV will make up a significant portion of overall power production as well. Though all REmap strategies are deployed, solar thermal heat can produce almost ten times more electricity than today's building and industrial sectors [13].

There is enormous potential for the usage of ML for power companies and investors to deploy more reliable and profitable processes that improve return on investment (ROI) and help the energy transition [15]. Although there are still various implementations of ML, it can speed up the phase of positive transition in the energy industry and business reducing carbon emissions that is becoming a bigger priority today. The United States Department of Energy (DOE) collaborated with International Business Machines (IBM) to create a Watt-Sun software, which measures vast quantities of weather data obtained from a large range of data sources and sites [16].

Another key advantage of moving 100% towards clean, renewable energy is the energy transition process will generate additional 11 million jobs in the energy sector by 2050 (see Fig. 3). The transformation will massively improve overall jobs in the energy sector by proactive policies. The change to green energy sources would produce more energy employment than the fossil fuel industry would lose. The renewable energy map case will lead to a loss of 7.4 million workers from fossil fuels by 2050, with a net increase of 11.6 million jobs generated by 19 million new jobs for renewables, energy conservation, grid improvement, and energy flexibility. Education and training policies would require the expertise and value-creating capabilities of these industries to fulfill the needs for the human capital of renewable energies and energy-efficient industries in the rapid growth process. A change that would lead to just and fair social and economic effects will prevent resistance. One of the most significant possible gains is to transform the socioeconomic environment.

Countries are just starting to realize the promise that ML brings and are integrating them into their policies. For example, China (adopting a nationwide smart grid using the current integrated technologies in AI), Australia (spending $25 million on ML and heading out to focus on energy in comparison to other countries), the United Kingdom (spending $119 million into the ML and robotics development programs in extreme

[1] Renewables: RISE Regulatory Environment, which countries have an enabling environment for investment in renewable energy? 25 Jun 2017, https://www.seforall.org/data-stories/renewables-rise-regulatory-environment. Accessed: 03/12/2021.

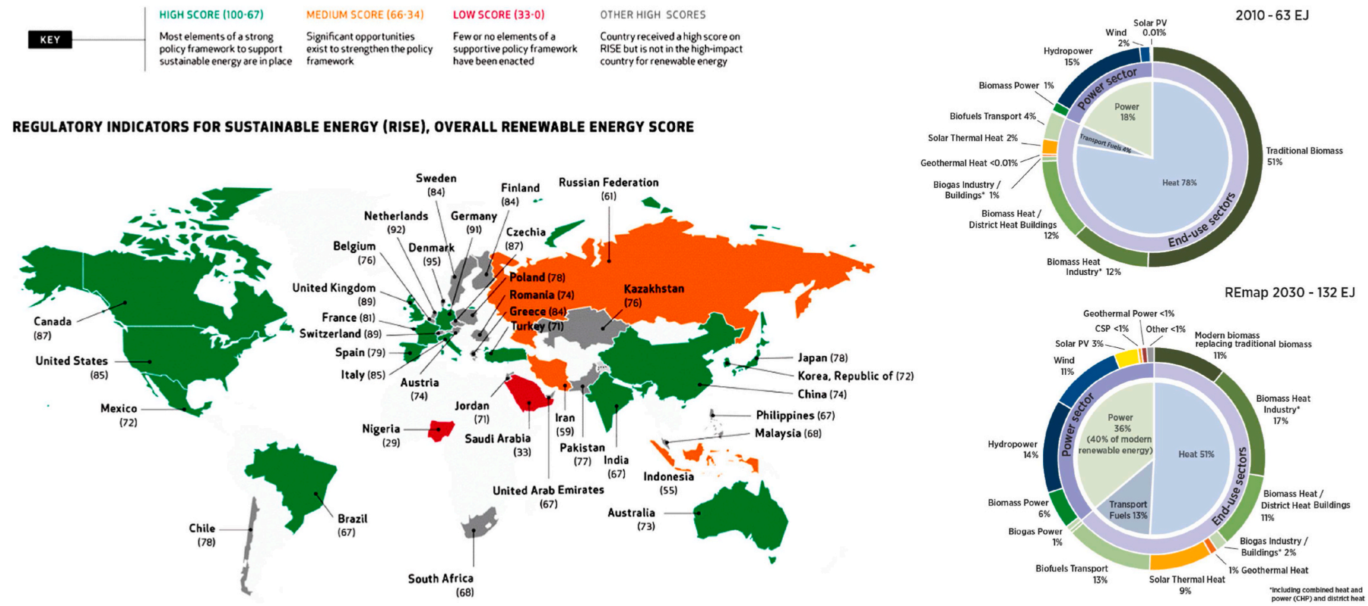


Fig. 2. An overview of global renewable energy usage in 2010 and the renewable energy map 2030 projections are divided into technologies and sectors.

climates, for application on ML and robotics innovations that are widely utilized in different conditions like offshore gas and oil fields), Italy (establishing 41 million smart meters for a project that began in 2001.

Consequently, Italy seems to have the most developed smart grid in the country), and the United State (the energy department proposed $30 million in support for ML) spending much to incorporate the ML technologies in the energy industry.[2] Also, large energy companies (e.g., ABB, Shell, Exxon Mobil, British Petroleum (BP), VIA, Stem, Pranav, Schneider Electric, Google, and DeepMind, etc.) are focusing on ML models. For example, integration between an ML technique and a major hydroelectric utility resulted in a 10% reduction in regular maintenance costs and an improvement of 2% in ABB company output.[2] BP company used an amazon cloud storage system was built with lubricant enterprise planning applications, resulting in a 40% quicker response time.[2]

1.2. Introduction to machine learning

ML is a subset of artificial intelligence (AI) that applies mathematical methods to data to learn the machine how to enhance its performance and accuracy without the support of external guidance [18]. ML focuses on applications that benefit from their previous experiences and improve forecasting and decision-making over time. We use probabilistic ML for two reasons. Firstly, it's the most successful approach to making decisions better under uncertainty. Secondly, probabilistic modeling is the main language for other scientific and engineering fields, facilitating its utility as a unifying framework [18].

How does ML work? To construct a ML program, there are generally four following steps:

1) *Prepare and select the training dataset:* Training data is a compilation of data that accurately represents the ML algorithm used to solve the issue it was developed to address.
2) *Select a method to be applied to the training set:* Algorithm classes vary depending on whether they operate with unlabeled or labeled data and on whether they solve a problem that is unlabeled or labeled.
3) *Training the algorithms to build the model:* Training the algorithms is an iterative procedure. First, we set variables to some values, then we execute the algorithms, compare the outcome with the desired results it could have generated, and make changes to the algorithm's weights and prejudices as needed, repeating the process before the algorithm returns the right or accurate answer most of the time.
4) *Improving and using the model:* Finally, the method is used with updated information and, if possible, to see if it becomes more accurate and more effective over time. New data will come from wherever the problem to be solved calls for it.

The ML models fall into four main classes, including i) supervised ML; ii) unsupervised ML; iii) semi-supervised ML, and iv) reinforcement ML. Table 1 shows a detailed overview and class functions of probabilistic ML models in energy distribution. Further, detailed analysis of different machine models, main categories/subcategories, and real-time implementation with codes are given in Ref. [19].

1.3. Historical overview of ML models

According to Fig. 4, the first stage started in the mid-1950s and continued until the early 1960s. Before that, the idea of ML was developed, with the primary focus on computer translation of logical thought. There was rapid development in symbolic theory, information engineering, and expert systems that arose as essential fields of study. Since 1960, ML study had phases of success accompanied by periods of stagnation. After completing the concrete evidence of all theorems in the book Principia Mathematica, Herbert Simon and Allen Newel developed the Logic Theorist, an automated theorem proving device. People slowly discovered.

That rational thinking capacity was not enough to achieve computer intelligence. Consequently, the analysis of AI inevitably reached its second level [20]. The second stage started in the 1970s and continued until the 1980s [20]. Scholars began to consider whether the information could be gathered and converted to a computer machine. In the 1990s, the third phase started. The ML discipline arose to cope with the information acquisition bottlenecks. Specifically, in 2006, ML joined the age of cognitive intelligence, which focuses on collecting big data, modern theoretical techniques, more efficient machines, and autonomous learning. These latest AI generations, such as cross-border integration, deep learning, SI, autonomous control, and human-computer collaboration, all carry out new features. The success of mathematical and computational learning theory has led to ML models' constant

[2] Machine Learning and Renewable Energy, https://www.pangea-si.com/machine-learning-and-renewable-energy/, Accessed: 03/12/2021.

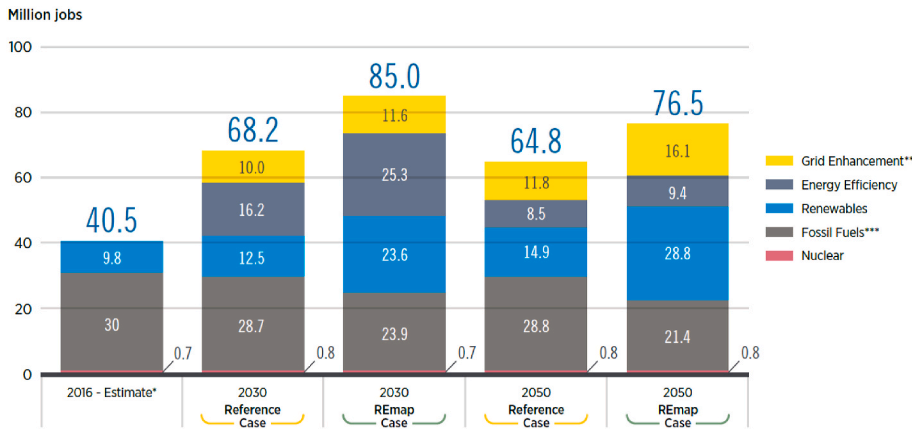


Fig. 3. The energy transition process will generate additional 11 million jobs in the energy sector by 2050 (*For 2016, there are no employment estimates for grid enhancement and energy efficiency. **Grid enhancement jobs apply to transmission and distribution grids and energy flexibility jobs, provided in the operation, development, and maintenance of infrastructure to allow the renewable energy integration of renewable energy sources into the grid. ***Both employment in the fossil fuel sector, including mining, refining, and consumption, are included) [17].

availability; representative ML techniques include deep learning, reinforcement learning, decision trees, and support vector machine (SVM)—the history of ML from 1957 to today as seen in Fig. 4 (right side). In the energy industry, the word ML is commonly used in conjunction with AI and is significant. On the other side, ML and AI are two separate entities because ML covers a portion or part of AI, but not all of it.

1.4. List of machine learning models and their usage in smart energy system

Several ML models listed in Table 1 are vastly used in smart energy systems. Here we will briefly elaborate on each model in detail.

Supervised ML, Unsupervised, and Reinforcement Machine learning: In the context of ML techniques, these terms are brought up with different definitions according to their use in different areas.

Supervised learning: It has got much attention in image classification, speech recognition, and machine translation. Power systems and smart energy systems play one of the key roles in data and load forecasting by providing input/output pairs to the ML algorithm. The ML algorithms do a function mapping on those pairs of input to their desired output which is also termed as a regression process. The key example is supervised vector machines (SVM) learning models used to analyze the electricity data generated by distributed generation and conventional generators [21].

Unsupervised learning only requires feeding input into the ML algorithm without any corresponding output (no dependency on output). Here, the clustering-based technique clusters the data into similar groups to hunt for hidden data groups or patterns of data. Grid-based clustering for energy systems can be incorporated for optimal energy distribution from the nodes. This energy clustering ensures the fast processing of unsupervised electricity data [22].

Reinforcement learning RL: RL doesn't operate over a fixed data set and has gained considerable interest in the energy domain. In the smart energy domain, some studies have compared the RL approach with the Model predictive control (MPC) method for uncertainty and forecasting; however, increasing complexities in the power system due to integration of renewable energies, has resulted in power system security along with bringing more uncertainties. RL did great in dealing with these smart energies based problems due to its sequential decision-making power taking account of risk measures and uncertainties [23].

Bayes Concept ML learning: Bayes theorem is a concept in the applied ML which provides the relationship between data and model. Bayesian methods can quantify that uncertainty, and deep learning models follow the Bayesian paradigm. In terms of smart power systems, the Bayesian methods could be used to predict net load forecasting. The aggregated net load forecasting in the smart energy power systems can be achieved by combining Bayesian theory and deep (LSTM) long short-term memory. LSTM is a recurrent neural network specially designed containing memory cells that can keep valuable information for the long term. It shows greater efficiency in dealing with load based on long-term dependencies, high volatility, and uncertainties [24,25]. Fully Bayesian inference on the other hand can be used to model selection choosing both predictive and evidence frameworks. In this regard, results from some studies have proved that the evidence framework outperforms the predictive method, which shows data overfitting [26].

The probabilistic ML models for smart energy systems also include Binary and Bernoulli distributions, Multinomial and Categorical distributions, and Univariate Gaussian distribution as described in Table 1. *The binomial distribution* is a probability distribution that summarizes the likelihood for a particular value among one or two independent values under a given set of parameters. *Binary and Bernoulli distribution* models have played a key role in the probability distribution of the smart energy power system. Researchers have adopted them for the Plug-in Electric vehicle PEV model to determine the probability distribution of EV charging patterns at different times of use [27,28]. Recently, most studies have adopted grey Bernoulli models to reduce randomness and uncertainty in the energy systems; hence, making predictions requires fewer operational data and analysis, especially in predicting the long-term development trend over data [29].

Logistic Regression: One kind of supervised ML used to model conditional probability. It is a univariate model, but it can be multiclass logistic regression. At the multi-stage decision process, the decision tree method has been commonly adopted to make a binary decision at each stage. Logistic regression, binomial and multivariate regressions brought different applications for smart energy system prediction such as power black outage prediction, load forecasting, etc. [30].

Common Univariate Distribution (CMD): CMD in ML is usually discussed separately by probability studies. Some of them are Student t distribution, Beta distribution, Cauchy distribution, Gamma distribution, Laplace distribution. *The beta distribution* is often proposed in PV irradiance and Wind speed calculation and forecasting by calculating its alpha and beta distribution [31]. *Cauchy distribution* is widely used in power system harmonic assessment, prediction-based models, wind power uncertainty calculation, along with the real-time dispatch of wind farms [32]. *Weibull and Gamma distribution* are also some of the methods commonly used for wind speed calculation in distributed generations (DGs).

Optimization Algorithms: Power system, in general, offers a variety of optimization technique for heterogeneous problems such as non-linear, large-scale, subject to uncertainties, discrete and continuous. There are certain optimization methods used in the modern power system which are addressed and mentioned by our study in Table 1. Those are; First-order and Second-order methods, Stochastic gradient descent, Constrained optimization, proximal Gradient method, Bound optimization, and Derivative and blackbox free optimization. The first-order optimization method is often used to categorize numerical optimization

Table 1
An overview and classes functions of probabilistic ML models in energy distribution.

Types of ML in energy distribution								
Supervised learning			Unsupervised learning			Reinforcement learning		
Classification based supervised learning	Regression-based supervised learning	Generalization and overfitting	Clustering-based unsupervised learning	Self-supervised learning	Evaluating unsupervised learning	Value-based reinforcement learning	Policy-based reinforcement learning	Model-based reinforcement learning
Probabilistic inference								
Bayes' rule		Bayesian concept learning		Bayesian ML				
Inverse issues	The Monty Hall issues	Discrete concept learning	Continuous concept learning	Fully Bayesian approach	Scalar input and binary output	Plug-in approximation	Binary input and scalar output	Scaling up
Probabilistic models								
Binomial and Bernoulli distributions		Multinomial and Categorical distributions			Univariate Gaussian distribution			
Sigmoid logistic function	Binary logistic regression	Log-sum-exp trick	Softmax function	Multiclass logistic regression	Cumulative distribution function	Regression	Probability density function	Half-normal
Common univariate distributions					Multivariate Gaussian distribution		Mixture models	
Student *t* distribution	Beta distribution	Cauchy distribution	Gamma distribution	Laplace distribution	Mahalanobis distance	Conditionals and marginals of an MVN	Gaussian mixture models	Gaussian scale mixtures, and mixtures of Bernoullis
Optimization algorithms								
First-order methods				Second-order methods				
Descent direction	Momentum methods	Step size and learning rate	Convergence rates	Natural gradient descent	Trust region methods	Newton's method	Quasi-Newton and other Broyden–Fletcher–Goldfarb–Shanno methods	
Stochastic gradient descent			Constrained optimization					
SGD for fitting linear regression	Iterate averaging	Preconditioned SGD	Lagrange multipliers	Linear programming	Mixed-integer linear programming	The KKT conditions	Quadratic programming	–
Proximal gradient method			Bound optimization			Derivative and blackbox free optimization		
Projected gradient descent	Proximal operator for L1-norm regularize	Proximal operator for quantization	The EM algorithm	EM for a GMM	EM for an MVN with missing data	Simulated annealing optimization	Grid search and random search optimization	Model-based blackbox optimization
Information theory								
Entropy						Relative entropy		
Entropy for discrete random parameters	Joint entropy	Cross entropy	Conditional entropy	Perplexity	Differential entropy for continuous random parameters	Forward vs. reverse KL	KL divergence between two Gaussians	KL divergence and MLE, and Non-negativity of KL
Bayesian statistics								
Conjugate priors					Non-informative priors			
The beta-binomial algorithm	The multivariate Gaussian-Gaussian algorithm	The Dirichlet-multinomial algorithm	Beyond conjugate priors	The Gaussian-Gaussian algorithm	Jeffreys priors	Reference priors	Invariant priors	–
Hierarchical priors		Empirical priors		Bayesian model comparison				
Hierarchical binomial algorithm	Hierarchical Gaussian algorithm	Hierarchical binomial algorithm	Hierarchical Gaussian algorithm	Bayesian model selection	Occam's razor, and Bayesian hypothesis testing	Posterior predictive checks	Bayes model averaging	Bayesian hypothesis testing
Approximate inference techniques					–			
Laplace approximation	Markov Chain Monte Carlo approximation	Online inference using assumed density filtering	Grid approximation	Variational approximation				
Bayesian decision theory								
Bayesian decision theory				Bandit problems				
Classification problems	Precision-recall curves and regression problems	ROC curves	Probabilistic prediction problems	Contextual bandits	Exploration-exploitation tradeoff	Optimal solution, and regret	Markov decision processes	Upper confidence bounds, Thompson sampling, and simple heuristics
Linear discriminant analysis and logistic regression								
Gaussian discriminant analysis					Binary logistic regression			
Quadratic decision boundaries	Nearest centroid classifier and	Logistic regression		Fisher's linear discriminant analysis			Nonlinear classifiers, and MAP estimation	Iteratively reweighted least squares

(*continued on next page*)

Table 1 (*continued*)

Types of ML in energy distribution								
	discriminative vs. generative classifiers		Linear decision boundaries, Naive Bayes classifiers		Linear classifiers, and maximum likelihood estimation	Perceptron algorithm, and stochastic gradient descent		
Multinomial logistic regression				Bayesian logistic regression				
Nonlinear and linear classifier, and hierarchical classification	Gradient-based optimization, and bound optimization	Maximum entropy classifiers	Robust logistic regression, and Bi-tempered loss	Laplace approximation	Approximating the posterior predictive	Variational inference	MCMC approximation	Online inference using assumed density filtering
Linear regression								
Standard linear regression	Ridge regression	Robust regression			Lasso regression	Bayesian linear regression		
		Student distribution	Laplace distribution	Huber loss		Estimating $\rho(w\uparrow,\ \sigma^2)$ with Gaussian prior	Estimating $\rho(w,\sigma^2 \uparrow D)$ with Gaussian-Gamma prior	Sparsity-promoting priors, Hierarchical priors, and empirical bayes
Generalized linear models								
The exponential family		Generalized linear models	Probit regression					
MLE for the exponential family, and exponential dispersion family	Log partition function is cumulant generating function		Multinomial probit models	Bayesian inference	Latent variable interpretation	Ordinal probit regression	Maximum likelihood estimation	–
Deep neural networks								
Multilayer perceptrons	Backpropagation		Training neural networks		Regularization		Feedforward networks	
	Forwards pass, and computation graphs	Backwards pass, and automatic differentiation	training the learning rate, variable initialization	Exploding and banishing gradients, and batch normalization	Early stopping, and dropout	Bayesian neural networks, and Sparse DNNs	Mixtures of experts	Radial basis function networks
Neural networks for electric images								
Adversarial neural networks					Solving different discriminative vision tasks with CNNs		Neural network basics	
Whitebox gradient-based attacks	Defenses based on robust optimization	Blackbox gradient-free attacks	Deep dream and neural style transfer	Real-world adversarial attacks	Objective detection and image tagging	Human pose estimation and image segmentation	Convolution in 1d and 2d, and convolution as matrix-vector multiplication	Boundary strides and conditions, pooling layers, and normalization layers
Neural networks for sequences								
Recurrent neural networks			1d CNNs	Transformers		Efficient transformers		
Vec2Seq sequence generation	Seq2Vec sequence classification	Backpropagation through time, gating, and long-term memory	1d CNNs for sequence classification and sequence generation	Self-attention, positional encoding, and multi-headed attention	RNNs and CNNs, Comparing transformers	Low-rank and kernel methods, and learnable sparse attention patterns	Recurrence and memory methods	Fixed non-learnable localized attention patterns
Nonparametric models								
K nearest neighbor classification	Learning distance metrics				Kernel density estimation			
	Convex, and linear methods	Deep metric learning, and classification losses	Ranking losses, and different training tricks for DML	Speeding up ranking loss optimization	Kernel regression	Density kernels	Parzen window density estimator	KNN and KDE classification
Kernel methods								
Mercer kernels		Gaussian processes		Scaling GPs to large datasets		Support vector machines		
Mercer's theorem, Hilbert space, representer theorem	Some popular Mercer kernels, Kernel ridge	Noise-free and noisy observations	GPs for classification	Random feature approximation, and exploiting parallelization	Sparse variational inference	SVMs for regression, Sparse vector machines	Multi-class classification with SVMs, relevance vector machines	Kernel ridge regression, and Converting SVM outputs into probabilities
Forests, trees, boosting, and bagging								
Regression and classification tree	Ensemble learning		Bagging tree	Random forests	Boosting tree			Interpreting tree ensembles with feature
	Stacking ensemble	Bayes mode averaging			LogitBoost, and gradient boosting		Forward stagewise additive modeling	

(*continued on next page*)

Table 1 (*continued*)

Types of ML in energy distribution								
						AdaBoost and exponential loss, and least squares boosting		importance and partial dependency
Learning with fewer labeled examples								
Transfer learning		Model-agnostic meta-learning and few-shot learning with matching networks	Semi-supervised learning				Active learning	
Fine-tuning supervised pre-tuning	Domain adaptation and unsupervised pre-training		Self-training and pseudo-labeling	Label propagation and consistency regularization	Deep generative models, and a combination of semi-supervised and self-supervised learning	Entropy minimization	Decision-theoretic approach, and batch active learning	Information-theoretic approach
Dimensionality reduction								
Principal components analysis	Factor analysis			Autoencoders		Manifold learning		
	EM model for PPCA and FA, and exponential family factor analysis	Generative model, Mixtures of factor analyzers	Nonlinear factor analysis and factor analysis models for paired data	Bottleneck autoencoders, contractive autoencoders	Denoising autoencoders, sparse autoencoders, and Variational autoencoders	The manifold hypothesis, and local linear embedding	Isomap, Kernel PCA, and maximum variance unfolding	Multi-dimensional scaling, Laplacian eigenmaps, and t-SNE
Clustering								
Hierarchical agglomerative clustering	K means clustering		Clustering using mixture models		Spectral clustering		Biclustering	
	The K-means++ algorithm, and the K-medoids algorithm	Vector quantization	Mixtures of Bernoullis	Mixtures of Gaussians	Laplacians encode the clustering for eigenvectors	Normalized cuts	Nested partition models	Basic biclustering
Recommender systems								
Explicit feedback			Implicit feedback			Leveraging side information	Exploration-exploitation tradeoff	–
Collaborating filtering	Autoencoders	Matrix factorization	Bayesian personalized ranking	Factorization machines	Neural matrix factorization			
Graph embeddings								
Shallow graph embeddings				Graph neural networks				Deep graph embeddings
Distance-based: Euclidean methods	Distance-based: non-Euclidean methods, supervised embeddings	Unsupervised embeddings and outer product-based: Skip-gram methods	Outer product-based: matrix factorization methods	Spectral Graph convolutions	Message passing GNNs	Non-Euclidean graph convolutions	Spatial graph convolutions	Unsupervised and semi-supervised embeddings

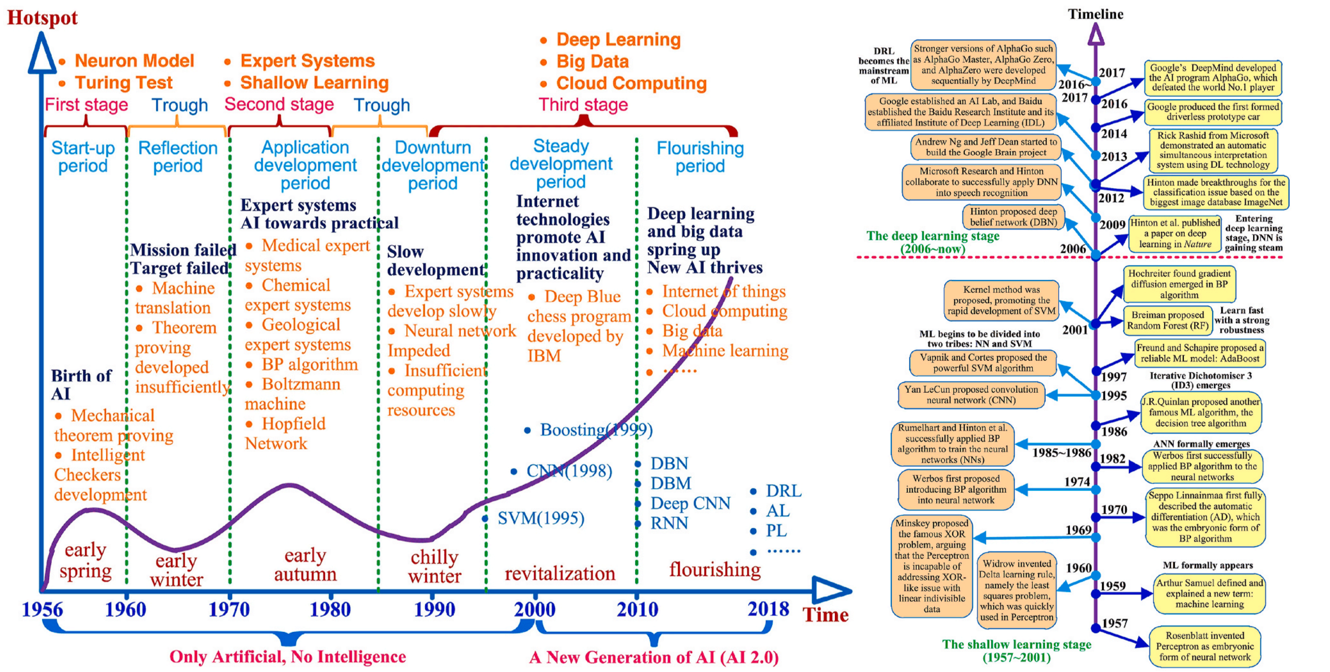


Fig. 4. A new AI generation: a technology perspective for ML application to smart energy and control systems and the evolution of ML from 1957 to today [20].

methods using a first derivative algorithm. The second-order method also called the Newton method is an algorithm that uses the second derivative in a scalar case. These optimizations play a vital role in the optimal power flow (OPF) problem of the energy power system. OPF is an optimization tool used for power system operation, analysis, and energy management. Linear programming, Mixed-integer linear programming, KKT conditions, and Quadratic programming are all bringing the application to energy and smart grid system in a variety of ways that include but are not limited to generation and expansion planning of power systems, operation of different kinds of conventional and modern power system, optimal power flow modeling and heuristic techniques, demand response and risk, uncertainty measures [33,34].

Bayes decision theory uses the Bayes theorem to find the conditional probabilities. It is based on tradeoff quantification among various decisions based on Bayes theorem probability and the cost associated with the probability. It deals with classification problems, ROC curves, probability predictions, Markov decision process, Thompson samplings, etc. In modern energy systems, including various kinds of renewable energy resources, the degree of uncertainty emerges as a huge problem. Bayesian decision theories stand out to effectively tackle the degree of uncertainties [35].

Deep Neural Networks: In the Smart grid and energy systems, the energy load forecasting and uncertainty brought by distributed generations remains a discussing topic for researchers. Deep neural network (DNN) consists of many ML-based methods to handle energy load and uncertainty problems differently. Convolutional Neural Network (CNN), Artificial Neural Network (ANN), and some other methods (refer to Table 1) have done a great job in dealing with the mentioned problems in modern energy systems [36].

1.5. How is ML changing energy distribution?

In reality, renewable energy utilities (e.g., solar, wind, nuclear, and hydro) have significantly gained from the recent increase in powerful AI, ML predictive models, and data science [37]. The predictive models have reduced their expenses, better forecast potential occurrences, and maximized their portfolio's return. ML models perform the different kinds of tasks in energy applications, for example, a) data storing [38], b) visualizing the data [39], c) data analysis [40], d) capturing data [41], e) data updating, and so on. Many reasons can be attributed to the growing interest in ML: 1) better computation power [42]; ii) hybrid transactional/analytical processing systems and big data [43], and iii) advanced and new ML techniques [15]. *How can ML help the energy industry?* Technology developments have enabled us to leverage the following approach in the energy industry: for example, i) ML and grid management [44]; ii) solving demand response [45,46]; iii) predictive maintenance [47]; iv) energy source exploration [48]; v) save costs when reducing electricity use [49]; and so on. Table 2 shows ML's high benefits and low complexity use case in core energy distribution function and real-time examples.

Pervasive sensing empowers extensive data, making it easier to utilize ML-related approaches [52]. In this scenario, statistical techniques-based ML algorithms may rapidly identify failure sensors and extract their data from analytics. Fig. 5 (a) illustrates a primary grid connected with microgrids (also known as "dispatchable electricity resources"). Energy production plants (renewable, conventional, etc.), power transmission lines, and power distribution networks are constantly regulated by utilizing cutting-edge technology such as ML models. Using the base stations in urban areas, sensor data can be transmitted via cellular communication networks. In certain rural places, satellites or LoRa could transmit sensor data.

Further, the sensor data is sent to the supervisory control and data acquisition system so that it can remotely track, automate the whole network, preserve network reliability, create alerts, and evaluate network topology. For billing purposes, the metering in residential, commercial, and manufacturing regions sends electricity usage and production data to the utility [53]. Additional data is frequently required to control the highly intermittent output of renewable energy. This information is forwarded to the demand response unit to better manage demand and supply. The processing and storage of sensor data are important for sensors to gather and report their findings to analytics tools. Fig. 5 (b) demonstrates the deployment of the sensors on distribution lines [54], transmission lines, transformers, underground cables [55], smart meters, circuit breakers, protective relays [56], etc., in a substation. In addition, a detailed analysis of pervasive sensing, ML in energy systems, and cyber-enabled grids is presented by Philip et al. [52].

Table 2
ML high benefits and low complexity use case in core energy distribution function: real-time examples.

Sr. #	Area of implementation of ML algorithms	Real-time applications at the glance
1	Yield optimization	GE renewable energy uses ML to develop on a cloud-based platform virtual wind farms that imitates an actual, physical design. This model works on a real wind trend and measures energy to maximize output on a single turbine scale or an individual level. It projects a 20% growth in electricity efficiency to save $100 million over a 100 MW farm's lifespan [50].
2	Energy forecasting	U.S.-based Vermont Electric Power Company is developing hyperlocal weather prediction systems through integrated data science and ML techniques. In both solar and wind farms in Vermont, the application of the weather model resulted in a decreased average solar energy prediction error by 6% and wind errors by 9%. Mr. Kerrick Johnson, VELCO's Vice-President of Policy and Marketing, said, we will save $1 million in fee payments per 1% reduction in load through improved resource orchestration. For example, as a transmission service in the short-term load prediction, we also recognize operational advantages [51].
3	Energy storage	Greensmith Energy uses ML to control energy storage devices and wider energy environments, a multinational energy storage firm. It supplies the Spanish island Graciosa, for example, with its real-time applications. With an incorporated 6 MW/3.2 MWh of energy store control facility, a modern 'Graciosa Hybrid Renewable Power Plant' can deliver 1 MW of solar and 4.5 MW of wind power to the local power grid, minimize the island's fossil fuel demand and reduce emissions of greenhouse gases substantially [51].
4	Grid behavior interface	The "Gridsense" technology is a Swiss leading energy corporation that uses ML to evaluate criteria such as grid requirements, demand, and generation of electricity, environment predictions, and pricing for electricity. This makes it possible for consumers to learn better. Using this information, GridSense optimizes customer and generator power use and decreases power grid peak loads, balance loads, and stabilizes power delivery [51].
5	Energy trading	In energy trading, British Petroleum (BP) uses ML automation. ML automation consolidates trading floor data utilizing automation robotics to imitate repeated procedures, CIO for enforcement, legal, risk, and finance at British Petroleum energy supply & trading. It changes the position of analysts and liberates them up more freely to concentrate on more important tasks. So, they can spend time analyzing and asking about the context of these data instead of collecting data sets [51].
6	Complaints management	Exelon created a channel-agnostic chatbot operated by ML, a US-based power and gas utility to address consumer problems on challenges like billing and outages. Exelon reduced the number of consumers and gained more insight into their consumption needs [51].

The implementation of ML algorithms includes yield optimization, energy forecasting, energy storage, energy trading, grid behavior interface, and customer complaints management. Detailed analysis of each approach is given in Section 2 and Section 3.

2. Machine learning in core energy technologies

ML is creating new opportunities for cutting-edge research in energy distribution [57]. This section covers seven different core areas of energy distribution, including i) advanced energy materials; ii) energy systems and energy storage devices; iii) energy efficiency in the industry; iv) energy demand response; v) strategic energy planning under uncertainty; vi) large-scale integration of renewable energy; and vii) big data analytics in the smart grid environment. Table 3 explains the ML technologies in core functions and technical use-cases in energy distribution systems [51].

This visualizes four significant aspects: power transmission and distribution use cases in core areas, energy services and supply, generation, exploration, production core use cases, and energy trading use cases in different energy distribution systems. Detailed analysis of each section is listed below:

2.1. ML for advanced energy materials

Recent advances in ML have raised the likelihood that data-driven materials research can revolutionize breakthroughs and offer new paradigms for producing energy materials. We can anticipate substantial improvements in materials science [7]. Moreover, recent developments in data-driven materials engineering often propose that ML technologies can improve the design and production of advanced energy materials in addition to the discovery and implementation of such materials. For illustrate, the National Institute of Standards and Technology (NIST) maintains 65 separate databases with 67,500 different measurements. Furthermore, during the last decade, more than 1.7 million research studies have been published on both solar cells and batteries [58].

Computational techniques are being established to help create structures and measure different electronic properties and other features [69]. Another noteworthy example is the materials initiative, which utilizes supercomputers to simulate all characteristics of the materials. Currently, there are more than 700,000 materials with forecast properties identified [60]. In 2016, Organization for Economic Co-operation and Development member countries spent over $16.6 billion on energy material research and development activities, while in 2000 was less than $10 billion [60]. In October, the United Kingdom unveiled its renewable development plan, which would spend more than $3.3 billion on low-carbon material technologies between 2015 and 2021 [58]. However, the enormous ability to transform this form of data into industrial and business applications is yet to be recognized [60]. ML algorithms designed to discover trends in large data sets — can significantly accelerate energy-materials discoveries. However, there are always challenges. Materials cannot be described universally. Specific physical properties, such as the elements' composition, shape, and conductivity, are essential for numerous applications. An abundance of scientific evidence on materials is hard to come by, and the feasibility of computational testing of theories relies on predictions and simulations that could be unreliable under the actual situation.

Different ML models for energy materials modeling are visualized in Fig. 6 [61]. The first key step in every ML research is to assess the objectives and forecast goals of the ML algorithms, which is normally achieved with the assistance of domain experts. This is arguably the most critical step since the goal must be unambiguously specified and theoretically learnable from available knowledge, such as composition and crystal/molecule structure, material information, determining quantities or objects experimentally, etc. In materials engineering, supervised learning seems to be the most general method of ML, in which a model has trained the mapping function between input feature variables (i.e., composition and molecular/crystal) and output labels or values (i.e., characteristics like bandgaps, energies, and so on). The parameters/weights of the algorithm are iteratively modified during training to mitigate model losses in response to training results. Both ML models aim to make low-cost, relatively reliable predictions that can be used

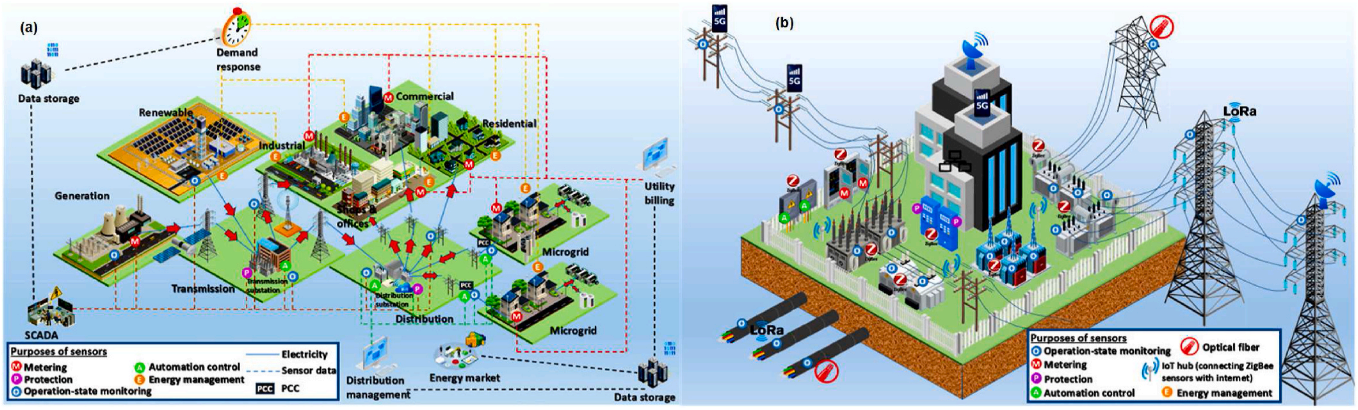


Fig. 5. Pervasive sensing (a) in the power grid and (b) in the substation side [52].

Table 3
ML technologies in core functions and technical use-cases in energy distribution.

Sr. #	Description	Power transmission and distribution use cases in core areas	Energy services and supply	Generation, exploration, and production core use cases	Energy trading core use cases
1	Use of AI in natural language	30%	34%	31%	27%
2	Use of AI in machine learning	14%	14%	16%	6%
3	Use of AI in computer vision	33%	24%	30%	36%
Intelligent automation advantages					
Sr.	Key areas of smart automation				
#	Description	Smart automation benefits as predicted as well as expected	Overestimated intelligent automation advantages in the energy sector	Underestimated intelligent automation advantages in the energy sector	–
1	Advantages of energy cost savings	14%	39%	47%	–
2	Consumer satisfaction advantages	16%	36%	48%	–
3	Incremental and new revenue advantages	17%	37%	45%	–
The energy and utility sector are driving significantly with the use of intelligent automation					
Sr. #	Sector by category	Improved consumer experience through faster response	Decrease the different number of steps and processes for purchase and queries	Increase the total number of consumer availability by being open longer hours	Personalized products/ services for consumers
1	Energy and utilities	81%	78%	74%	67%
2	All sectors	60%	61%	30%	48%
Compared to other sectors, the energy sector brings huge benefits from smart automation					
	Description	Increase in energy operations quality	Data accuracy improvements	Improvement in workforce agility	Increase productivity in staff members
1	Energy and utilities	40%	37%	33%	32%
2	All sectors	30%	30%	20%	26%
	Description	Turnaround faster time for service requests	Consistency in data improvements	Less requirement of resources to accomplish the designed task	Accurate and better compliance with regulatory
3	Energy and utilities	32%	32%	28%	27%
4	All sectors	21%	20%	18%	21%

instead of more costly computational, experimental, or human-driven approaches [62,63]. As a consequence, ML models make it possible to:

1) Providing fast forecasts of characteristics and novel materials to speed up the discovery of new materials.
2) At greater time/length scales, precise simulations of complex materials are feasible.
3) Further, it enhances the interpretation and characterization.

Collaboration between energy sciences and ML communities is required. One of the keys aims of the energy and renewable energy materials research task, administered by the mission innovation global partnership, is to help for implementing the ML approaches. Mutual government commitments support this partnership, and governments must include adequate funding to meet their commitments. In summary, a substantial increase in investment is expected in ML-enabled materials science worldwide. Experimentalists, ML researchers, and algorithm programmers can work together to help with troubleshooting more quickly [64]. The training samples or, more generally, computed—are obtained in the second step. Featurization is converting experimental data into numerical properties (usually in the form of tensors or vectors) that can be used to differentiate between various materials [7].

2.2. ML toward advanced energy systems and storage devices

A more comprehensive range of storage technologies has been built to address the daily energy needs of the utility grid. It includes a battery (e.g., flow batteries, capacitors, advanced chemistry batteries, and a

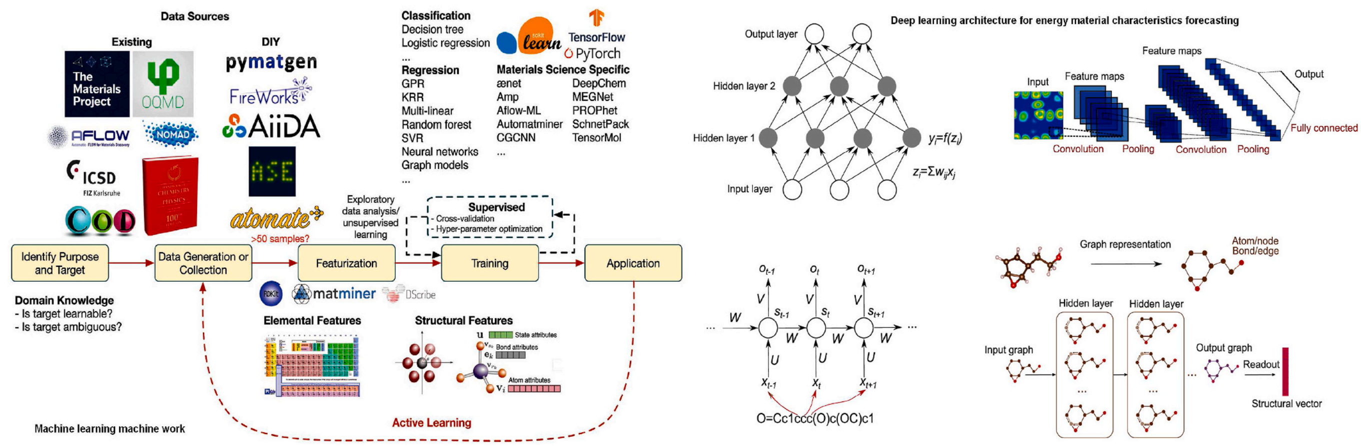


Fig. 6. A framework for developing ML models for energy materials. This workflow consists of five major steps: defining an objective, data processing, data quality improvement, model creation, and implementation [61].

large number of electrochemical storage solutions) [65], hydrogen (e.g., excess energy production is converted into hydrogen via electrolysis and stored) [66], thermal (e.g., captured cold and heat to create on energy offset or demand needs) [67], mechanical storage (e.g., gravitational energy, harness kinetic energy or other innovative technologies) [68], and pumped hydropower (e.g., using water to build large-scale energy reservoirs) [70].

As technology advances, energy storage systems and devices with increased capacity, better efficiency, longer life, and a more intelligent management approach are anticipated [70]. Advanced control methods depend on certain indicators' instantaneous status, so developing those structures needs a trade-off among many parameters. ML will significantly speed up simulations, catch dynamic structures to boost predictive precision and make smarter decisions based on specific data information. Fig. 7(a) demonstrates the use of ML in different energy storage devices and systems. Primus Power,[3] a pioneer in low-cost, grid-scale electrical energy storage systems, has developed a flexible, stable distributed flow battery device that can be used for different storage applications (see Fig. 7 (b)). Fig. 7 (c) represents the overall distributed concepts of ML in advanced energy systems. It is divided into three main classes including: 1) energy domain [71]; 2) energy components [72]; and 3) energy applications [73,74]. Saqib et al. [75] have presented a detailed analysis of these classes. Continuing to grow the percentage of renewable energy providing electricity to the grid as part of the energy transition phase ensures that the amount of produced power can be impacted by conditions, making it harder to monitor. A smart grid must be built and combined with energy storage technologies to effectively control and dispatch electricity to enhance grid reliability and renewable energy usage. The bulk of this can be done by storing excess wind and solar resources in an energy storage facility. Such devices serve various industrial and corporate purposes, including peak shaving, contract capacity optimization, and backup control. This form of energy conservation is commonly recognized as the most adaptable green energy approach possible (see Fig. 7 (d)).

ML algorithms will benefit the energy storage industry greatly. ML-enabled energy storage can assist in the processing and reviewing data and provide insights into maximizing power consumption and forecasting future faults by simulations. Incorporating battery-based intelligent storage into a renewable energy system almost inevitably improves the economic benefit. ML-enabled storage can allow real-time flexible storage availability, resulting in improved demand for both the consumer and the grid. Since renewable energy sources are intermittent, ML may help address the problem of catching the intermittent variations of generation phases.

To conclude, the demand for ML in energy storage is now greater than ever [76]. Incorporating it with energy technologies would affect the planet's future. The convergence of ML and energy storage is a logical as well as a practical starting point. However, the system is only in its development stage in terms of the possible effect on the energy storage system. This synergy has the power to transform the environment and open up many possibilities while simultaneously enhancing sustainability.

2.3. ML tools for energy efficiency

ML's energy effect will rapidly become unsustainable unless new methods are created. But how can we allow the responsible use of it? Manufacturers rapidly utilize ML to increase throughput and reduce energy usage (see Table 4). The EU-funded FUDIPO initiative is making substantial progress in incorporating ML into various critical process sectors on a broad scale to gain significant energy and resource savings [77]. An example of the energy performance of active buildings contracting techniques toward resilience and smart buildings is visualized in Fig. 8. [3]As an electrification-focused on basic demand response to differential pricing is often the best approach to a thriving business scenario, ML can play a significant role as aggregator sites of implicit demand response in buildings. There is a lot of ML research that aims to improve predictive algorithm accuracy, but scientists and engineers are recently getting more active in improving energy efficiency.

Energy[4] use impacts millions of schools, households, hospitals, and industrial facilities, each imposing its energy conversion strategies [78]. Consumers in the US spend billions of dollars/years on energy conservation hoping that these investments would compensate for themselves by lower electricity prices in the future [80]. Many studies have been conducted on energy efficiency optimization algorithms so far. The most prominent categories include: i) simulation-based optimization methods [81], and ii) surrogate-based optimization methods [82]. In the simulation-based optimization method, the objective function (e.g., energy usage as well as thermal comfort) is controlled by an optimization method, and the decision parameters are controlled by an optimization tool to develop the objective function iteratively. There are

[3] Roberta D. Angiolella, Active building Energy Performance Contracting (AEPC) models towards smart flexible buildings, https://www.construction21.org/articles/h/active-building-energy-performance-contracting-aepc-models-towards-smart-flexible-buildings.html. (Accessed: 03/12/2021).

[4] Jim Kring, Grid-Scale Energy Storage Powered by LabVIEW, https://blog.jki.net/news/grid-scale-energy-storage-powered-by-labview, Accessed: 03/12/2021.

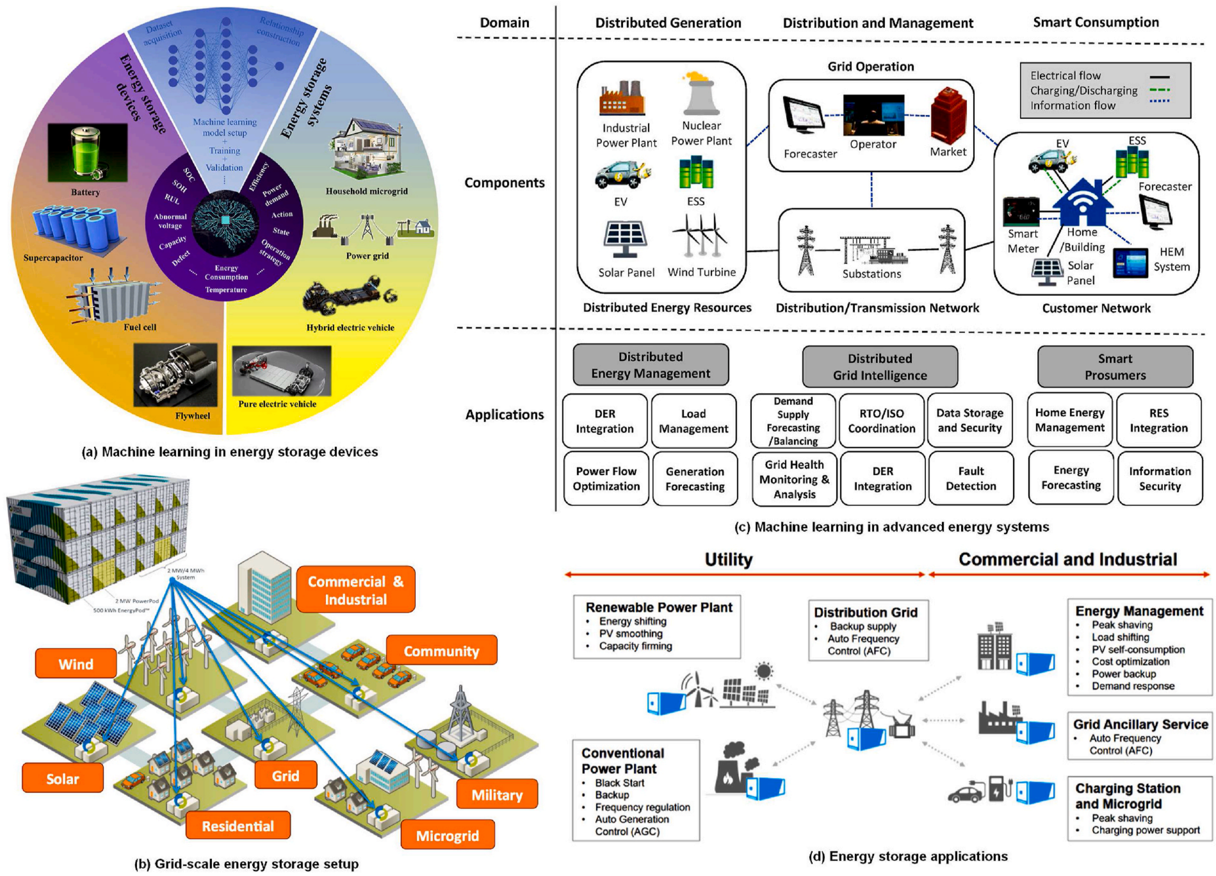
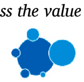
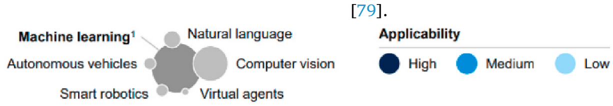


Fig. 7. (a) ML in energy storage devices [68]; (b) grid-scale energy storage solutions [75]; (c) use of ML in advance energy systems; (d) real-time energy storage applications.

Table 4
ML designed to help record huge improvements across the value chain in the energy industry [187].

Sr. #	Description	Applicable technologies	Produce	Project	Provide	Promote
ML can produce values across the value chain						
1	Retail		Automate the store operations and optimize the merchandising energy devices and products at micro space	Anticipate energy demand trends, however, automating and optimizing supplier contracting and negotiation	Tailored, enriched, and convenient customer experience	Services, and products at low prices, with the right target and the right message
2	Power utilities		Improve electricity generation yield, optimize preventive maintenance, prevent electricity theft, and decrease energy waste	Enhance energy supply and demand, measure the reliability of integrated energy production assets, and automate the control of energy demand-side response	Provide accurate energy consumption insights, automate the consumer services with the use of virtual agents	Day-of-day dynamic tariffing as well as match producers of different consumers in real-time
3	Manufacturing devices		Reduce the error at different states, improve the process tasks, reduce the energy material delivery time, and limit the different product rework	Enhance the device efficiency and yield, and automate supplier requirements and anticipate different parts of requirements	Enhance the skills for pilot training and maintenance engineer	Optimize the energy pricing, forecast of maintenance services, and refine sales-leads prioritization
The ML techniques can help capture a significant number of chains across the value chain						
1	Power utilities	–	20% energy production increased with the use of ML models and smart sensors to optimize the energy yields, improvement in EBIT 10–20% by applying ML to increase the predictive maintenance, automate the line fault forecasting and enhance the capital productivity [79].	The key aim of reducing the 10% in-country energy usage is by applying the deep learning models to forecast the energy supply and demand [79].	There is a huge saving in energy bills, for illustrating $10-$130 by applying ML models to automatically switch the energy supply deals [79].	–



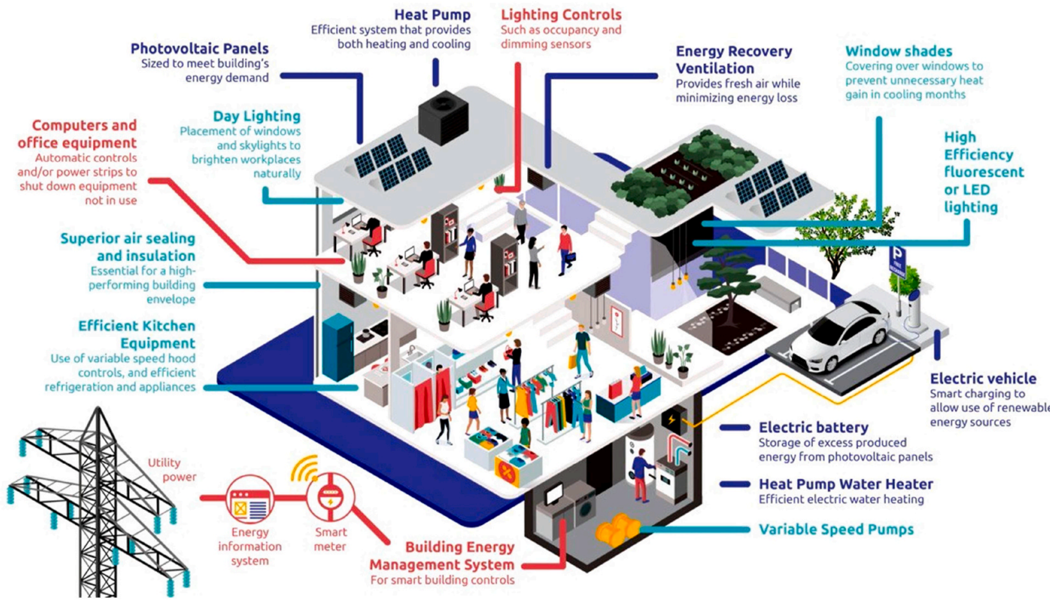


Fig. 8. Techniques for active building energy performance contracting in the direction of smart, flexible buildings.

different types of simulation-based optimization methods including: i) optimization methods [83]; ii) derivative-free methods; iii) gradient-based methods; iv) global search (meta-heuristic) methods; v) local search methods; vi) population-based methods; and vii) single solution-based methods. Surrogate-based optimization has a lot of potentials, particularly for high computational systems. However, the few experiments that have been performed so far have not looked at how to build surrogate models effectively, have neither they thoroughly exploited their benefits in terms of facilitating optimizing improvements. The usage of both surrogate-based and simulation-based optimization approaches in buildings is still a hot topic of research. Both approaches, though, have a high computational cost when searching for near-optimal solutions.[5]

According to the reference case (which considers existing and expected policies, like nationally determined contributions), energy-related carbon dioxide emissions will rise slightly per year until 2040, then drop slightly by 2050 to approximately match current levels (see Fig. 9). It is an improvement over the 2017 report, which reported that annual carbon dioxide emissions would be higher in 2050, showing that nationally determined contributions and the steadily improving cost and efficiency of renewable energies affect long-term energy planning and scenarios. Severe additional cuts, furthermore, are required. Yearly energy-related carbon dioxide emissions would also decrease by 2050 from 35 gigatons (in the comparison case) to 9.7 gigatons, reducing even more than 70%, to reach a climate target of restricting global warming to 2 °C [17].

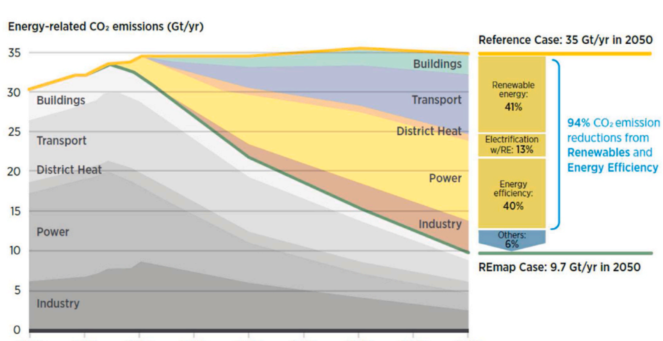


Fig. 9. Energy efficiency and renewable energy can render over 90% of the decrease related to energy carbon dioxide emissions [17].

According to international renewable energy agencies' study, energy conservation and green energies, along with deep electrification of end-uses, will provide over 90% of the required reduction relating to energy carbon dioxide emissions. For particular production processes, pollution, fossil fuel switching (to natural gas) and carbon capture, and sequestration in manufacturing can be used to make up the difference. Nuclear power generation will remain unchanged from 2016. Simultaneously, by 2050, substantial efforts must be made to cut carbon emissions from industrial processes and land use to less than zero. Without success in those areas, the temperature target will not be met. ML can help in this regard. ML improve the efficiency of power system infrastructure, helping to integrate the renewable energy with the smart grid, improving the efficiency of conventional power plants through automation of power grid infrastructure, and enhancing the performance of buildings, transport, district heat, and industrial sector, which are the key part of energy distribution.

2.4. ML techniques for energy demand-side response

ML is one of the most important parameters for demand-side management and market response since it allows consumers to adjust the market mechanisms in actual environments, thanks to ML enabling technologies [84,85]. Smart homes are using digital technology to help with demand management. The ML techniques link devices, including rooftop solar PV, local battery storage, smart meters, and home appliances to the internet, allowing for data collection and sharing [86]. Fig. 10(a) demonstrates the synergies between developments for demand-side management integration [87]. It includes four essential parts, including i) enabling technologies (e.g., renewables power to heat of residential sector, batteries behind the meter, ML/AI, and big data applications; and electric vehicle for smart charging); ii) market design (e.g., net billing schemes and time-of-use tariffs); iii) business models (e.

[5] Delta Energy Storage System: A Versatile Power Modulation Strategy, https://blog.deltaww.com/en/energyinfrastructuresolutions-en/products/delta-energy-storage-system-a-versatile-power-modulation-strategy, Accessed: 03/12/2021.

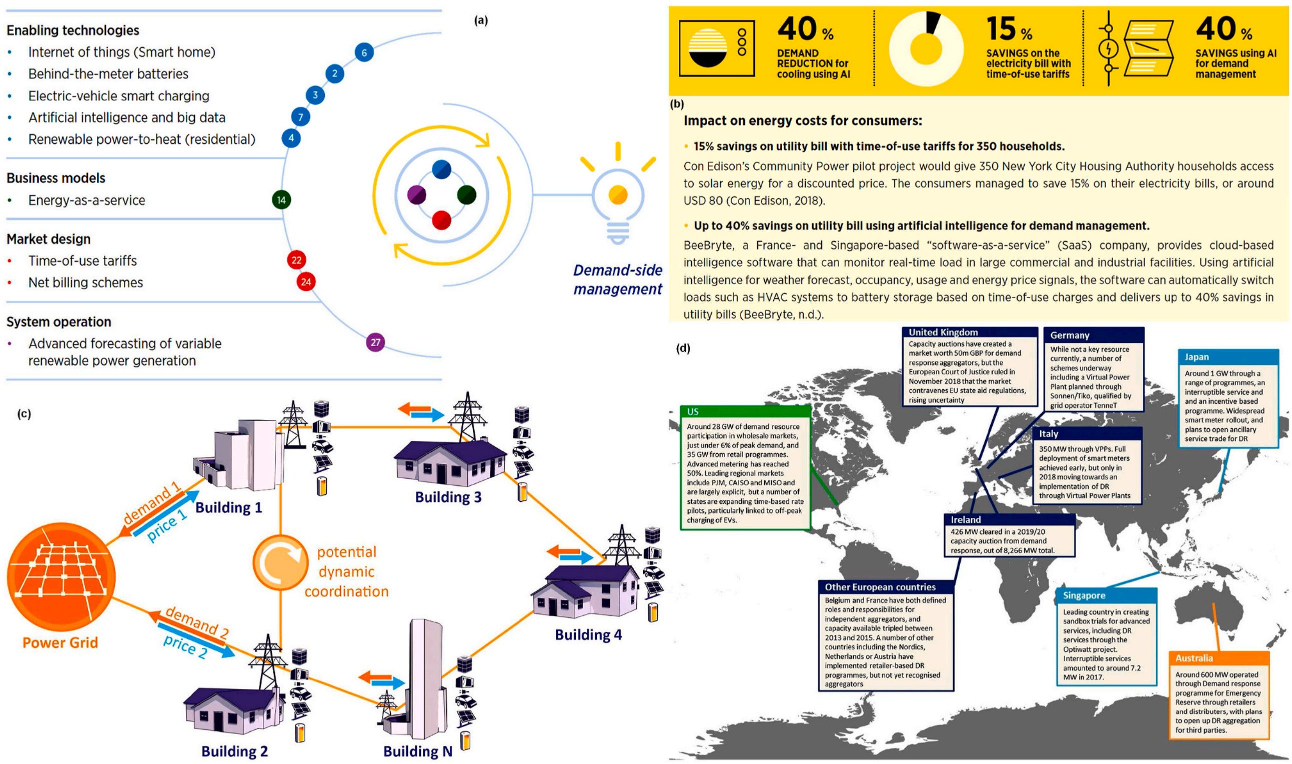


Fig. 10. ML for (a) synergies between developments for demand-side management integration [87]; (b) 40% energy demand reduction using AI [87]; (c) multi-agent based demand-response [88]; (d) in the sustainable development scenario, from 2018 to 2040, there is a huge potential for demand response.[7]

g., energy-as-a-service); and vi) system operations (e.g., renewable power production and advanced prediction of parameters).

According to a new study, the numbers of "connected homes" or "smart homes" increased by 31% from 17 million to 29 million in the three years between 2015 and 2017, reflecting a compound annual growth rate of 31% [87]. In a pilot project in Sweden that used market mechanisms and ML models for demand response, peak energy usage fell from 23% to 19% of overall electricity demand; 17% of peak demand was moved to non-peak hours.[6] A 1% decrease in power prices for a utility result in a 0.66% decrease in peak demand for that system on average.[6] According to the American Council for an Energy-Efficient Economic, demand-response programs are being used to reduce power demands by 10% or more. Fig. 10(b) explains the energy reduction for cooling using AI-enabled ML models. Con Edison's Community Power Pilot Project provides solar electricity at a reduced cost to 350 New York City Housing Authority households. Customers could save 15% on their utility bills, or about USD 80. A detailed analysis of these facts and figures using AI-enabled ML in demand is given in Ref. [87].

Fig. 10(c) demonstrates the demand response structure for multi-agent coordination [88]. Buildings can be seen exchanging information with the electricity grid, as their price and demand of electricity, while providing some kind (cooperative or competitive) dynamical communication and/or sharing of information between them. Fig. 10(d) visualizes the sustainable development scenario; from 2018 to 2040, there is a potential for demand response. Despite its tremendous potential, demand-side resilience success is erratic across the world.

Further, there are two main demand-response groups, including i) price-based energy demand response [89]; and ii) incentive-based energy demand response [90]. Price-based energy demand response includes real-time energy pricing, rates based on time of use, and critical-peak pricing. The incentive-based energy demand response includes emergency energy demand response programs, control of the direct load, ancillary market services programs, curtailable/interruptible rates, capacity market arrangements and programs, and buyback programs/demand bidding [91].

There are different kinds of demand response challenges including: i) advanced metering investment (e.g., meters that record consumer energy usage based on time-of-day or more regular billing blocks are required for all time-based rates); ii) customer desire/inertia for simplicity (e.g., most consumers (mainly residential consumers) would be suspicious of programs that involve initiative, such as when the program's proposed framework is not transparent); iii) reduction in demand (e.g., the ability to sell gaming participants can provide or state that they would reduce as they have already shut down energy requirements for the day is one with the main problems of most assessment approaches); iv) voluntary and mandatory participation price-based programs (e.g., experience shows that the degree of consumer engagement and overall load declines is moderate where the participation in price-based services were voluntary); v) fair time-based and simple pricing (e.g., customers informed of everyday pricing and market increases by different methods achieve better results and are comfortable with the programs); vi) peak load demand during off-peak hours (e.g., in the demand response), the key strategy is to change demand throughout peak load hours or high energy rates. There could be a scenario when a large number of consumers are likely to change their energy demand in response to the energy price, contributing the grid to reliability issues); and vii) the intermittent nature of renewable energy sources (e.g., the renewable energy sources are highly unpredictable and intermittent. Renewable energy sources (e.g., solar and wind) will improve their participation in profitable electricity markets by joining demand response by a virtual power plant) [91].

[6] IEA, Demand Response, https://www.iea.org/reports/demand-response, Accessed: 03/12/2021.

[7] IEA, Demand Response, https://www.iea.org/reports/demand-response, Accessed: 03/12/2021.

2.5. ML in smart manufacturing in the smart grid paradigm

The modern technological revolution is known as Industry 4.0 [92]. A large amount of data is produced by connecting any machine and operation to the internet using network sensors. ML is used to analyze produced valuable information and data about smart manufacturing. Different kinds of ML models are used in smart manufacturing in the smart grid paradigm. The most prominent models include: 1) Bayesian networks [93]; 2) k-nearest neighbors [94]; 3) SVM [95,96]; 4) instance-based learning [97]; 5) ensemble methods [98]; 6) artificial neural networks [99]; 7) multiple logistic regression [100]; 8) decision tree [101]; 9) gradient boosted [102]; 10) self-organizing map [103]; 11) k-means [104]; 12) bag of words; 13) random forest [105]; 14) locally weighted learning; and 15) additive models [106].

The development of smart manufacturing is linked to technical advancements and the changing needs of business models and stakeholders. Smart energy manufacturing systems collect real-time energy consumption data, improve decision-making accuracy, increase plant reliability and performance, and boost overall productivity [74]. Fig. 11 (a) depicts the development of smart manufacturing. ML is used to stimulate autonomous smart manufacturing schemes, in which key performance indicators and adaptive specifications are smart manufacturing, and input data framework is the key objective. The authors argue that conventional manufacturing processes in transition can understand specific criteria from four different viewpoints: objective, practical requirements, technical requirements, standards, and market requirements. Fig. 11(b) depicts the goals and conditions of smart manufacturing using ML [107]. The key objective of smart manufacturing in the energy sector includes i) autonomous lean operation (e.g., the primary goal of creating autonomous lean smart manufacturing is to improve the production system's performance and autonomy); ii) sustainable value added (e.g., the goal is primarily concerned with smart manufacturing's long-term value and viability of smart devices); and iii) win-win partnership (e.g., the ultimate goal of smart manufacturing in autonomous service environments is to produce continuous and better communication in which the whole smart

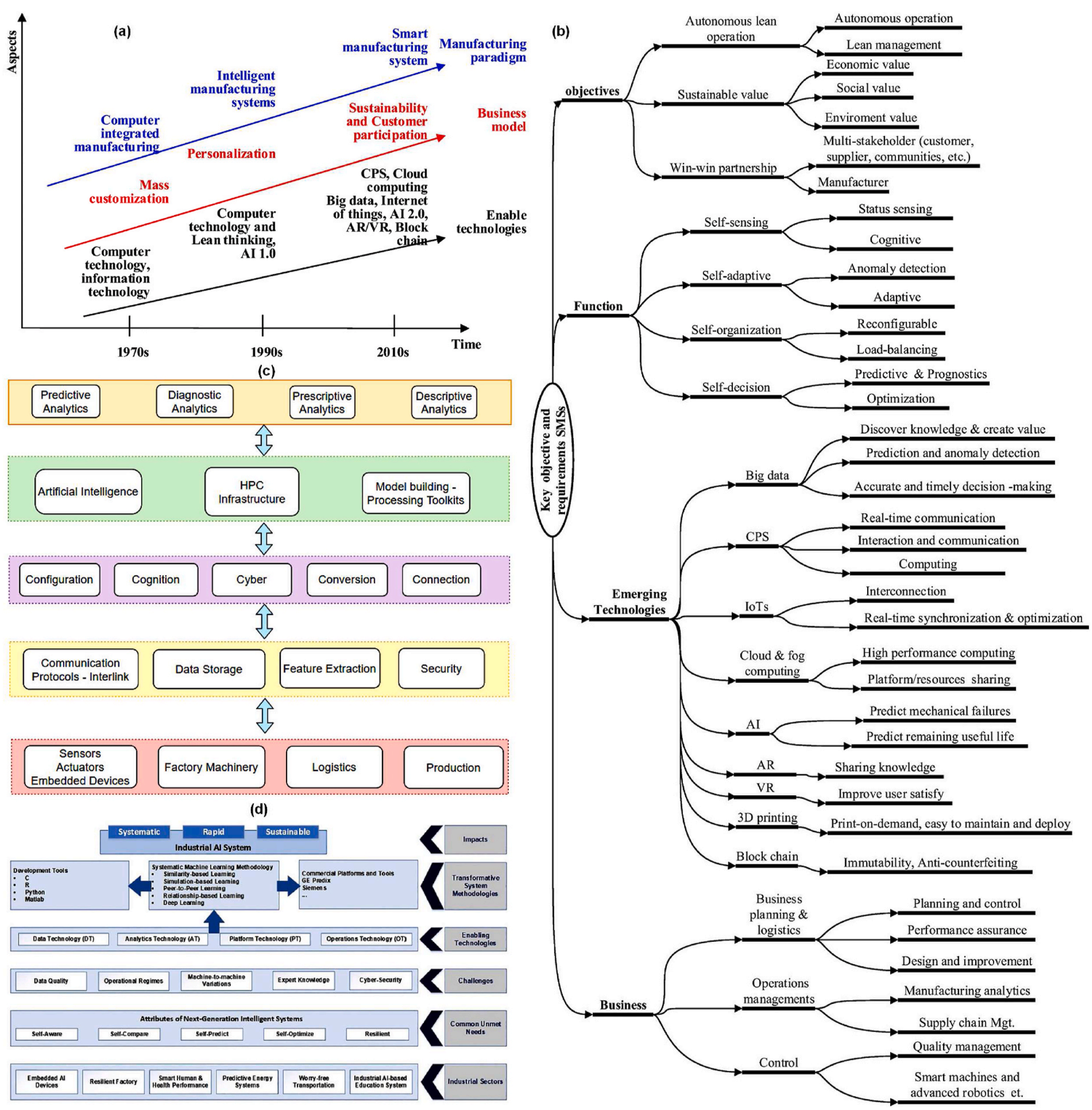


Fig. 11. ML in smart manufacturing (a) the advancement of intelligent manufacturing systems; (b) requirements and objectives of smart manufacturing [107]; (c) industrial AI architecture [106]; (d) industrial ML structure [106].

manufacturing lifecycle is more successful than the number of its individual elements). Smart manufacturing's ultimate goal in autonomous service environments are to produce continuous and better communication in which the whole smart manufacturing life cycle is more successful than the number of its different pieces or elements of different smart energy devices. No wonder integrating and bridging a variety of emerging technology helps smart manufacturing accomplish autonomous lean operations and long-term value-added. Industry 4.0 has three key dimensions: it promotes the computerization or digitization of manufacturing industries, reduces delivery times, provides greater reliability, adapts to consumer needs with limited batch sizes, and expands downstream service offerings.

Fig. 11(c) demonstrates the ML structure in the industrial sector. The first layer, which starts at the bottom and works its way up, involves gathering information from the different energy systems devices. Data from manufacturing equipment, data about the duration/summation/faults of output, and data knowledge about the logistics procedure are essential to building a well-performing and robust industrial ML framework. The different hardware components used to acquire the data, such as actuators, sensors, and embedded ML units, cannot be ignored. The next layer creates coordination links between industry sectors and ensures effective data acquisition. It is also in charge of storing the data gathered by the first tier, securing it with the right software and procedures, and extracting useful functionality from a preprocessing method if necessary. The industrial ML architecture in Ref. [108] as six levels of the pyramid shown in Fig. 11(d). The industrial sector is included in the first layer. The ML models, health, and human performance, the resilient factory, productive energy systems, industrial ML-based energy systems, and worry-free transportation are embedded ML devices. The second layer is designed to meet the demands of Industry 4.0. The third layer emphasizes the complications/challenges that an ML system could face. The fourth layer demonstrates how to construct industrial ML systems using enabling technologies. The programming methods, deep learning or ML methodology, and platforms used to build the framework comprise the fifth layer. Finally, the sixth layer refers to the potential effects of an AI framework. Thanasis et al. [106] have been conducted a detailed study on ML in innovative manufacturing industries.

2.6. ML for strategic energy planning under uncertainty

Decision-makers have long been interested in energy forecasting, management, and operation, especially regarding energy demand and resource allocation [109,110]. Researchers have adopted various decision-making techniques—and eventually implemented by managers—to assist decision-makers in selecting optimal choices, designing energy strategies, and supporting competitive alternatives to energy in uncertain settings [111]. Strategic energy planning aims to use the ML models to increase awareness and identify the overall market needs for incorporation into the basic and advanced energy infrastructure, as well as to reach an agreement on the best integration strategies and industry practices focused on these strategies (see Fig. 12)[8] [112]. Moreover, Table 5 shows the real-time implementation of ML models in different energy systems.

One of the key ML in the energy management industry drivers is AI's usage to improve the smart grid performance [106]. A grid is a network of cables, transformers, and different parts of infrastructure components that carry energy over long distances. It may be for a whole country or a transcontinental grid, which distributes electricity throughout continents and countries. The grid's power comes from several resources, including solar panel systems, fossil-fuel-based power plants, hydroelectricity plants, wind power plants, and nuclear power plants, making grid service challenging. They become energy-efficient and reliable as ML analyzes the large amounts of data obtained from their processes daily, elevating the ML in energy design management and market. For example, Table 6 demonstrates that ML transforms the customer and operations services using ML models. It is further shown that the new investment in distributed energy generation by 2040 would be $10.4 Trillion [113–115]. The total investment in consumer empowerment, grid modernization, and industrial size, and energy internet is noted 700 Million, $400 Billion, and $14 Trillion, respectively [113–115]. The ML in the energy management industry is expected to rise at a 19.8% compound annual growth rate from $4439.1 million in 2018 to $12, 200.9 million in 2024, up from $4439.1 million in2018.[9] In 2018, the utility group led the industry by the end customer, as ML technologies were introduced by several organizations, including Duke Energy Corporation and Dominion Energy Inc., to fill the difference between power demand and supply.

Incorporating the internet of things (IoT) in energy management is the most important development in the ML in the energy industry. IoT enables remote access, program automation, data analytics, and intelligent tracking [116]. It employs smart meters and sensors mounted in both assembly lines and equipment to notify consumers of the energy used volume. Most utilities believe ML would have a significant effect on their businesses, but they have yet to incorporate it into their key strategies. The smart grid may be improved in terms of control and performance for energy performance strategy. ML forecast load demand and energy supply, enabling fully automated the world's largest thermal solar farm in California, for example, to avoid millions of dollars in damages due to mismatches between forecasted demand and real generation [117]. Furthermore, linking small, localized networks such as micro-grids to the cloud allows them to exchange and create data and usage, allowing them to forecast load better. In reducing energy management expenses, trading operations can be done automatically using ML-based algorithms [117].

The uncertainty of the modern energy landscape poses serious challenges for utilities and independent power producers. Traditional market models, which are marked by consistent supervision, vertical integration, guaranteed long-term return on investment, and non-intermittent generating developments, are being overwhelmed and outpaced. New innovations in ML that are affordable, autonomous, and digitized would take their position, requiring new business opportunities in the process.

2.7. ML support for large integration of large-scale renewable energy

According to a survey, renewable energy production must rise eight times larger than the fastest pace to reduce global warming or heating (see Table 7). The IREA claims $131tn will be required in investment in clean energy over three decades (see Table 8). The IREA said that immediate steps are needed to respond to increasing energy demand, which would entail the total expenditure of $131tn in renewables by2050[10]. The agency IREA anticipates that the global energy consumption would decrease to 4% by 2050, although the gas peak would fall to 6% in 2025 and coal would fall to 2% by the mid-century.[9] The increase in power consumption indicates that just over half of all energy used by 2050 will be renewable electricity, contrasted with 21% by 2018.[9] Approximately two-thirds of the energy consumed by fossil fuels have been consumed in recent decades but will be limited to 10% by 2050.[9]

[8] EPRI Smart Grid Demonstration Initiative, https://smartgrid.epri.com/doc/EPRI%20Smart%20Grid%20Demonstration%202-Year%20Update_final.pdf.

[9] AI in Energy Management Market, https://www.reportlinker.com/p05842955/AI-in-Energy-Management-Market.html?utm_source=GNW, Accessed: 03/12/2021.

[10] The Guardian, https://www.theguardian.com/environment/2021/mar/15/renewable-energy-growth-must-speed-up-to-meet-paris-goals-agency-says: Accessed: 03/12/2021.

Fig. 12. Smart grid energy resources and energy management.

ML can revolutionize renewable energy: To develop a significant rise in renewable energy's intermittent nature while maintaining the grid stability, ML would be needed in the low-carbon transition (see Fig. 13)[11] [37].

1. ML has the power to unlock the enormous green energy potential. If organizations do not accept it, it may risk falling behind the curve.
2. ML's strong analysis capability can boost market forecasting and resources management.
3. ML's automation capabilities will help organizations achieve operational success in various ways.

The energy market is confronted with major global challenges that need immediate action. Policy initiatives to a net-zero energy future, like the endorsement and implementation of the Paris Agreement, necessitate a rapid transition to a low-carbon economy [118]. As policymakers scale up renewables and move dependence on fossil fuels, major disruptions to the energy market are anticipated [119]. Grid operators, engineers, and customers are trying to harness ML as we step into the 4th Industrial Revolution, opening the way for a seamless shift to more green energy [92]. The capacity of ML to have better prediction tools enables enhanced asset management and market forecasting, although the automation technology drives operational excellence, resulting in a strategic edge and increased efficiency for stakeholders.

How do ML techniques transform renewable energy? ML could improve several aspects of business, from market forecasting to asset maintenance. Forecasting power levels has become critical to ensuring a reliable and productive system when growing megawatts are fed into the utility grid from intermittent variable renewable energy resources. This is attributed to the lack of baseload production from different sources like coal, which renders power grid inertia by high concentration spinning machineries like gas and steam turbines, while renewables take up a larger share of the grid. Power grids would be unreliable and vulnerable to blackouts if system inertia is not present. Wind and solar generation will now have a massive amount of wind and solar power production data, enabling ML to forecast different capacity levels, thanks to the implementation of sensor technologies [120]. More reliable forecasts of intermittent renewable energy over shorter timescales enable utilities and energy investors to predict their performance better and compete in the balancing and wholesale markets–all while avoiding penalties. In the meantime, ML algorithms based on massive volumes of weather data will help grid operators make the most use of their power system networks and utility power grids by adjusting swift power system operations to change the climate circumstances/conditions at any time [76]. Improved short-term renewable energy prediction will lead to greater unit engagement and dispatch.

performance, improving flexibility and reducing the need for operational reserves. According to Brian-Case, a Chief Digital Officer at General Electric Renewable Energy, unexpected disturbances in the energy industry will cost 3.01–8.01% of capability and USD 10 billion in the annual lost-generation cost.[12] ML-based algorithms analyze industrial data to render machine health forecasts and suggest measures to increase asset quality, such as wind and solar farms.

The challenges of implementing ML through the industry: The poor nature of energy consumption data, customer regulatory, and mistrust challenges could all be obstacles for the new energy technologies. There is concern that is depending too heavily on ML might expose energy networks to cyber-attacks in today's information age. The hacking of 30 substations in Ukraine in 2015 served as a wake-up call, leaving 230,000 residents in the darkness for 6 h.[9] One year later of this incident, a quiet more minor attack on a transmitting station in Kiev happened. Audit, data bias, and continuing algorithm testing are problems ML frameworks must consider when planning implementations from a performance perspective [63]. Data must be extracted and rendered computer-readable, ensuring that accuracy is maintained from beginning to end. Frequent data checking is needed for trustworthy ML to ensure that techniques stay in effect over time and that devices do not deviate significantly from the original techniques as they learn. Skepticism has been raised that depending too enormously on ML might expose energy networks to cyber-attacks. Advanced ML automation

[11] Energy systems integration, University of Twente, https://www.utwente.nl/en/et/tfe/research-groups/TE/research/research/Energy_Systems_Integration/, Accessed: 03/12/2021.

[12] Brian Case, Chief Digital Officer at GE Renewable Energy, https://www.ey.com/en_dk/power-utilities/why-artificial-intelligence-is-a-game-changer-for-renewable-energy, Accessed:03/12/2021.

Table 5
ML for smart grid paradigm in energy distribution.

Sr. #	Description	Main areas ML play an integral role
1	Modeling and optimization	1) Google's Deepmind is the leading in the area of modeling and optimization in energy distribution, ii) reinforcement-learning algorithms are used for energy modeling, iii) probabilistic models used for solar irradiance and wind power forecasting, iv) IBM proposed sophisticated energy generation and weather forecasting in the United States from 15 min up to 30 days in advance.
2	Consumer-facing services	1) Integration of storage and microgeneration with virtual power plants, ii) smart home energy management, iii) ML used to identify customers' lifestyles, iv) social media and fintech, and v) storage or demand response, vi) virtual power plants (e.g., artificial neural networks to decrease supply forecasting errors).
3	Investment and markets	ML agents are largely used for investment models and market interactions.
4	Security and maintenance	i) ML used to forecast and optimize the maintenance efforts and schedules, ii) physical and cyber security measures to calculate the energy assets, iii) previous maintenance cycles, iv) failure modes, v) component lifetimes, vi) outages cost to optimize replacement and maintenance, timetables (e.g., used decision tree and artificial neural networks), vii) deep learning models are used to identify the energy consumption trend and flow, and viii) rendering early warning of disruption.
5	ML make the smart grid smarter and decrease the requirement for power utilities to add new power plants	i) Sensors and ML make by-minute changes to optimize production energy efficiency through adjustment to wind change, ii) ML-enabled forecasting anticipates the demand and supply peaks and increase the usages of different intermittent renewable energy sources, iii) smart wired combined with ML to enable the real-time power system dispatching as well as optimize the current building's asset portfolios and current grid loads, iv) insect size robots and drones used to forecast the power system failure and inspect the power system infrastructure without interrupting energy generation, v) ML and smart meter enables power utilities to offer different kinds of services which based climate conditions, energy usage, and other factors, vi) different virtual agents used to automate the call centers which is based on service history, and further ML offers different kind of early warning of bad debts, vii) fields workforce receive real-time updates as enable to decrease the response-time to decrease the large impact of outage.

systems, on the other hand, are disconnected from information technology structures and have no network links between them, making them far more challenging to penetrate. In summary, with advanced market forecasting, advanced asset control, and organizational excellence by automation, ML has revolutionized the renewable energy industry concerning time [121]. ML's capacity to incorporate a massive rise in variable renewable energy into a predictable and secure grid will be needed if the low-carbon transition accelerates.

Table 6
ML transforming the customer and operations services.

Category	% Improvements
Increased the operational efficiency	78%
Reduced false positives rate	68%
Reduced operational cost because of process improvement	75%
Greater regulatory/legal compliance at lower cost	70%
Reduced churn	66%
Enhanced customer satisfaction	73%
Reduced the customer complaints	72%
Japan got success enhancing the turbine efficiency and reducing the operating and maintenance cost	20%
McKinsey's UtilityX achieved replacement and maintenance cost savings	10–25%
Billing, metering, & security	9%
Decentralized energy trading	33%
Carbon trading and green certificate	7%
Energy gird management	8%
ML opportunity and revolution	
New investment on distributed energy generation by 2040	$10.4 Trillion
Consumer empowerment	700 Million
Grid modernization	$400 Billion
Industrial size and energy internet	$14 Trillion

2.8. ML for big data analytics in smart grid

ML helps power utilities in all big data cycle steps, including collecting, preserving, and retrieving different forms of data from various sources [122]. These include pattern, data, content, action, decision, risk, and goal management [123]. ML algorithms improve the potential of IoT platforms and big data analytics to add value to both of these business segments. There are three categories of IoT data, according to the author: (1) raw information or unstructured data [84], (2) metadata [43], and (3) valued-added or transformed data [124]. ML can help assist three of these data categories to categorize, define, and make decisions. By 2024, the global demand for ML in IoT and big data would surpass $24 billion. By 2024, embedded ML in the service of IoT and objects would be worth $4.6 billion globally.[13] Asia Pacific would lead the overall demand for IoT, ML, and AI in big data, followed by North America. By 2024, ML in industrial machines would be worth $415 million, with collaborative robots growing at a 42.5% compound annual growth rate. ML will be a critical AI technology in realizing the full potential of IoT, big data, specifically in edge computing platforms.

Fig. 14(a) represents the three-basic structure of sources of data in the power grid, including 1) structured data [7]; 2) unstructured data [125]; and 3) semi-structured data [126]. The dimensions and format structure varied in terms of many data sources in the smart grid. The data include the distribution stations and electrical information from distribution switch stations, non-electrical information like marketing, smart energy meters, and regional economic and meteorological data. The processing and collection of these data are critical for power subsystem operation, plant scheduling, marketing business conduct, and critical control equipment repair. These kinds of data come from different energy sources, for illustrate, smart metering infrastructure, micro-phasor determining unit, phasor measurement users, remote terminal units, smart appliances, geospatial network topology, climate

[13] Research and Markets, Artificial Intelligence in Big Data Analytics and IoT Markets, 2019–2024: Focus on Data Capture, Information and Decision Support Services, https://www.globenewswire.com/news-release/2019/05/21/1833672/0/en/Artificial-Intelligence-in-Big-Data-Analytics-and-IoT-Markets-2019-2024-Focus-on-Data-Capture-Information-and-Decision-Support-Services.html, Accessed: 03/12/2021.

Table 7
Renewable energy map 2030, an overview of energy transaction indicators for deploying AI-enabled ML technologies [188].

Description	Unit	2000	2012	Renewable energy map 2020	Renewable energy map 2030	Renewable energy map reference (%)	Reference case 2030	Compound annual growth rate 2012–2030 (%/y)	Compound annual growth rate 2000–2012 (%/y)	Indicator for renewable energy map 2030
Technology indicators										
Pumped hydro	GWe	689	1004	1350	1600	06	1508	0.6	3.2	–
Hydropower excluding pumped storage	GWe	689	150	225	325	06	306	4.4	N/A	–
Wind offshore	GWe	17	06	50	231	242	68	22.5	N/A	300000 of 5 MWe wind plants
Wind onshore	GWe	17	283	600	1404	56	900	9.3	26.4	
Concentered solar power	GWe	0	03	15	83	62	52	12.5	7.6	830 of 100 MWe solar plants
Ocean	GWe	–	01	03	09	519	02	17.3	–	–
Solar photovoltaics	GWe	08	100	400	1250	184	441	15.1	23.5	12.50 million of 100 kWe solar plants
Electric vehicles	Million	–	0.20	25	160	133	69	45.80	N/A	10% of the total passenger car fleet
Biomass, traditional	EJ/year	28	27	20	12	−58	29	−4.3	−0.0	–
Battery storage	GWe	N/A	2.0	25	150	105	73	27.1	N/A	5% of total parameter renewables capacity
Biomass, advanced for cooking	EJ/year	–	01	04	04	88	02	8.4	10.4	270 million 5 kW_{th} cookstoves
Number of heat pumps	Million	N/A	04	15	40	58	25	13.3	N/A	–
Biomass heat from cogen	EJ/year	01	03	04	14	129	06	9.8	10.2	–
Heat pump	GW_{th}	N/A	50	177	474	58	300	13.3	N/A	–
Biomass pellets for heating	EJ/year	0.1	01	02	03	49	02	5.80	48.6	16 million 20 kW_{th} household boilers
Geothermal heat	EJ/year	0.2	0.5	0.7	1.2	86	0.6	4.3	9.6	–
Biomass pellets for heating	EJ/year	0.1	01	02	03	49	02	5.8	48.6	16 million 20 kW_{th} household boilers
Share in industry	%	–	01	09	99	968	03	41.8	–	–
Biomass chips logs for heating buildings	EJ/year	–	05	05	06	49	04	1.0	6.4	31 million 20 kW_{th} household boilers
Share in buildings	%	100	99	91	67	−31	97	10.50	−31	–
Biomass boilers industry including biogas	EJ/year	04	04	05	07	0	07	3.4	−1.0	0.7 million 1 MW_{th} industrial boilers
Solar thermal area	Million m^2	157	446	1162	4029	163	1532	13.0	11.3	–
Biofuels transport	Billion liters/ year	18	105	214	650	127	287	10.7	15.9	15% of worldwide transport fuel consumption
Total biomass use	EJ/year	43	51	61	108	37	79	4.3	1.4	20% of total primary energy supply
Financial indicators										
Fossil fuel subsidies	USD billion/ year	–	544	–	–	–	–	–	–	–
Net incremental system cost	USD billion/ year	–	–	–	133	–	0.9%	–	–	–
Subsidies need	USD billion/ year	–	101	–	265	–	58%	–	–	–
Net incremental investment requirements	USD billion/ year	–	–	–	265	–	1.7%	–	–	–
Regional indicators based on renewable energy map 2030										
Global – modern + access	(%)	–	–	–	30	–	–	–	–	–
	(%)	–	09	–	27	–	14	–	–	–

(continued on next page)

Table 7 (*continued*)

Description	Unit	2000	2012	Renewable energy map 2020	Renewable energy map 2030	Renewable energy map reference (%)	Reference case 2030	Compound annual growth rate 2012–2030 (%/y)	Compound annual growth rate 2000–2012 (%/y)	Indicator for renewable energy map 2030
Worldwide – modern renewable energy excluding traditional biomass										
Global – modern + access + energy efficiency	(%)	–	–	–	34	–	–			–
Global – modern + access + energy efficiency + "renewable energy+"	(%)	–	–	–	>36	–	–	–	–	–

Table 8
The road map of future energy generation and production.

Generation source	Modern energy storage and conversion technologies		Nature of load
Solar power, Wind power, Biomass, Hydropower, and Nuclear Power	Photovoltaic panels Photochemical conversion Wind turbines Photochemical conversion Hydropower Nuclear power	Batteries flow Alkaline Ion batteries CO2 storage and capture Fuel cells Hydrogen production Capacitors	Building load, industries load, and transportation load (e.g., electric buses and vehicles)
Key supporting technologies for electric power and smart energy systems			
Smart grid			
Energy internet	Commercial users	Artificial intelligence	Analyzing
Energy services producers	Household users	Cloud computing	Optimizing
Energy traders	Industrial users	Big data	Judging
Energy production	Smart buildings	IoT	Decision-making
Smart electric power and energy systems	–	–	–

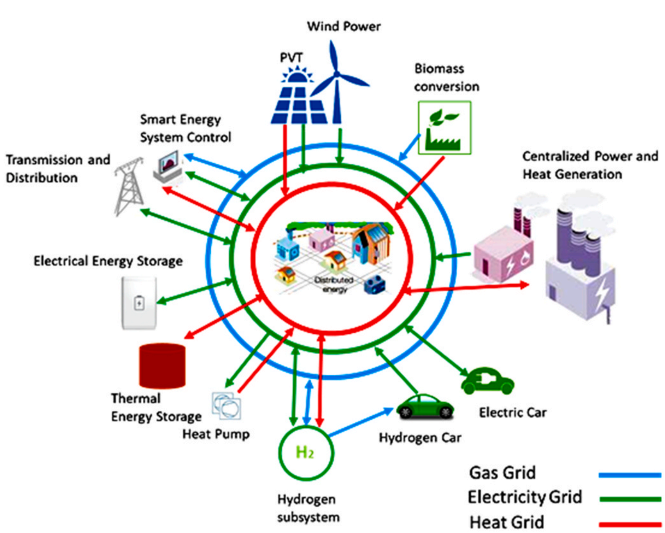


Fig. 13. Renewable energy integration.

information, social media, traffic information, data acquisition, and supervisory control, asset inventory, smart plugs, network switches, transformers, programmable thermostats, and field measurement devices as visualized in Fig. 14(b). Five different characteristics of big data in the smart grid environment, including 1) value; 2) volume; 3) veracity; 4) velocity; and 5) variety, are shown in Fig. 14(c) [127]. As a result, big data analytics would be essential for the effective operation of potential utility power grids and for implementing appropriate business strategies for key stakeholders (i.e., system operators, electric utilities, aggregators, and consumers).

Data analytics can include descriptive, predictive, informative, and prescriptive analytics, depending on the possible usage cases. As seen in Fig. 14, descriptive approaches are often used to characterize grid and consumer organizational behaviors, while diagnostic models examine working conditions and grid operator decisions (see Fig. 14(d)). The diagnostic paradigm focuses on determining a case's triggers, making it appropriate for proactive action. Forecasting is often required to anticipate operational environments and potential decisions since data analytics' primary goal are to include a preventive approach [130]. Individual functional and devices units produce thousands of terabytes of data each year, as visualized in Fig. 14(e). Due to the vast amount units of this type (i.e., sensors, customers, substation), and utility grid operations (i.e., home energy control, management of power distribution, distributed energy resources DERs, and management), power utilities must maintain millions of terabytes of climate and energy consumption data, which is expected to grow in the future. As a result, utilities must examine what increasing data entails for their conventional activities and devise methods to extract value from large data quantities. According to a global study of industry respondents and 1000 electric utility from ten countries, big data analytics is critical for a potential smart grid and a source of potential market prospects for the majority (80%) of electric utilities [129]. The Utility Analytics Institute predicts that data-related costs will continue to decline, as seen in Fig. 14(f). Every 14 months or so during the last 30 years, the expense of storing data has also been reduced. For example, the storage of a gigabyte of data costs around $11,200 was noted in 1995, $11 in 2000, and three cents today [131]. Real-time data gathering and distribution have become commercially viable due to declining data storage and maintenance costs, creating significant utility opportunities to develop profitable ML business models.

In addition, extracting and mining useful patterns from a large amount of data input for forecasting, long-run decision-making, and a different key to success is inference to big data. However, there are unique challenges that exist for ML and data analytics for illustrating different kinds of format variations in the raw data, streaming data that moves at a breakneck pace, the reliability of the predictive analytics, widely dispersed input sources, high dimensionality, poor and noisy

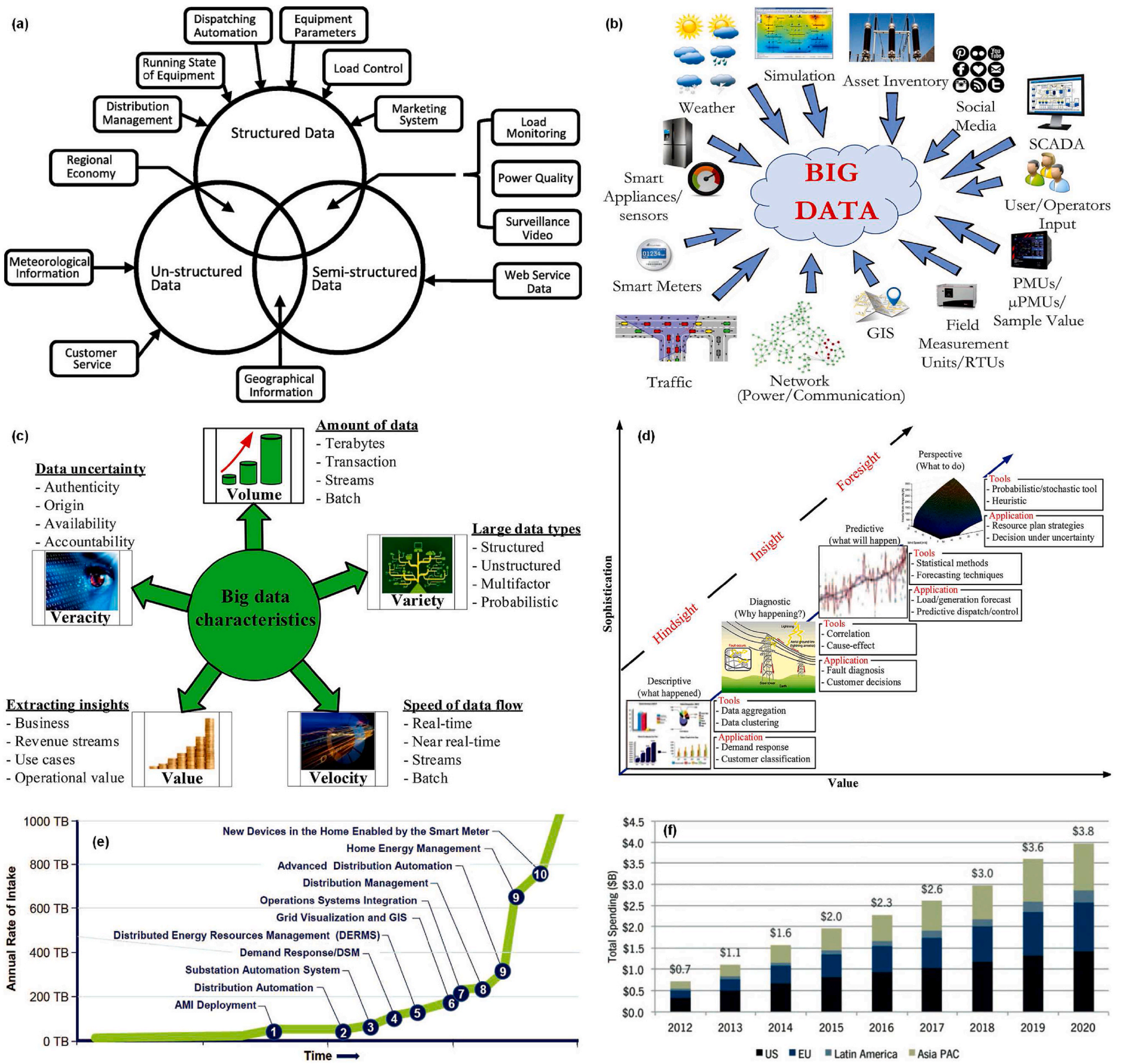


Fig. 14. (a) Sources of data in the power grid [128], (b) electrical and non-electrical sources big datasets in the smart grid environment, (c) big data key characteristics of smart grid, (d) an overview of various analytics approaches and their key applications in smart grid, (e) big data pattern volume in power utilities, and (f) world utilities spending expenditure on data analytics [129].

quality of data, methods scalability, imbalanced energy consumption and climate input data, un-categorized and unsupervised data, the limited number of labeled/supervised data, etc. Big data allows ML algorithms to discover finer-grained trends and render more timely and precise forecasts than ever before; however, it often poses significant challenges/obstacles to ML, like model scalability and distributed computation [132].

3. Use cases of ML for energy distribution utilities

The growth of renewables, the transition to a smarter grid, and marketing tactics transform the electricity generation environment and place stress on utilities' profit margins. On a scale, there is a greater need to make better decisions. To stay successful, these decisions must be taken quickly. ML quickly becomes the most powerful method for making smarter energy distribution decisions and building stronger consumer relationships. In this section, we will cover the use of ML models in different energy distribution sectors, and detail of each is listed below:

3.1. Accurately forecast the energy prices and consumption requirement

Consumers and companies gradually generate their electricity as individual power production (using wind and solar power) becomes more manageable and cheaper [15]. Users can generate, use, and store their energy by individual power production. Based on where they reside, they might also be willing to sell excess electricity back to the local power utility [133]. ML can help determine the most advantageous period to generate, store, or sell this electricity back to the utility grid. When rates are low, electricity should be used or retained/stored, and when rates are high, it can be sold out to the utility grid. We may render even more precise hourly estimates using ML algorithms to look at past data, energy consumption patterns (e.g., solar and wind), and weather forecasts. This helps individuals and businesses with electricity generation systems make rational choices on how to use their resources. The adaptive-based neural fuzzy inference approaches, for example, have been used to forecast short-term wind and solar forecasts for energy generation [124]. This encourages energy producers to cultivate energy supply while still selling it back to the grid at peak rates (see Table 9).

Any power company has to forecast the electricity needs of its consumers accurately [134]. To present, there is no appropriate energy storage solution, which ensures that the energy has to be distributed and used almost immediately after it is generated—the use of ML techniques to improve the predictive performance [135]. Historical data on electricity usage, weather predictions, and the kinds of companies or buildings working on a particular day all contribute to calculating the energy consumption. For instance, if it is a warm summer day in the middle of the week, office buildings will consume more electricity so that the air conditioning will be running at full capacity. Weather predictions and historical data will aid in the early detection of these trends, preventing rolling blackouts triggered by air conditioning units during the summer. ML identifies complex variations in a variety of driving variables (such as the day of the week, duration, forecast solar radiation, and wind power, past energy demand, major sporting events, air temperature, mean energy demand, pressure and moisture, and wind direction to understand the energy demand fluctuations. ML allows more precise decisions than humans since it can detect more energy complex patterns [136]. This means that when we consume electricity, we can maximize efficiency and lower prices without making costly changes or adjustments.

3.2. Accurately forecast the merit order of energy prices

Utility companies can choose from various energy sources, including alternatives such as wind and solar, as well as fossil fuels and nuclear power [137]. These various resources are sorted into a price-based merit order [138]. This establishes the direction in which the different sources of power are sold. ML can be used to evaluate both historical and real-time data due to access to information from a wide range of sources. ML algorithms are easier to account for all of the other various pricing factors – weather conditions, energy supply, demand, available resources from the particular sites, historical energy consumption, etc. – to foresee an optimized order of merit. This is particularly useful in markets with a lot of penetration of renewable energy sources, like solar and wind, since these sources' energy supply is difficult to predict.

Fig. 15 illustrates a stylized summary of sustainable energy production for 24 h. In the short term, of a one-day-ahead energy market, it is believed that energy consumption is inelastic. Since supply providers must acquire renewable energy supplies in advance, the residual production load that must be bought on the power markets is limited accordingly. As a result, the assured feed-in of energy produced by renewables reduces energy consumption. The German merit-order curve, which is a step function of single plant units in the real world, is simplified as a linear supply curve in the diagram. Reduced demand on the markets tends to lower costs as long as the slope of this supply curve is positive [139]. Since energy consumption and renewable electricity production change hourly, calculating the merit-order effect occurs real value is much more complicated than estimating the market value. Permanently falling electricity costs of production – particularly in renewable energy – have shifted the merit order chain, with traditional power plants falling further behind (see Fig. 16).[14] The result is obvious with the increased feed-in of renewable energy sources (i.e., photovoltaics, wind energy, or biomass).

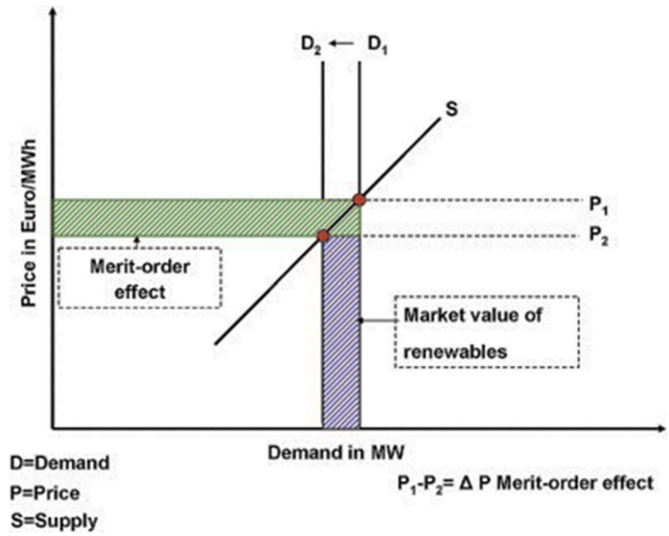


Fig. 15. The impact of merit order on renewable energy production [123].

During peak load times, varying wind, and PV power generation forms with different marginal costs near zero (0) value are moving values into the energy market, bringing traditional power production plants to the end of the merit order [140]. The merit order impact of clean energy is how the electricity sector relates to this trend [15]. Traditional power plants only meet the domestic load, which is the remaining energy requirement that renewable energy cannot meet.

The merit order rate is a steadfast representation algorithm that works very well for describing the short-term power industry development. Determining the long-term evolution of energy rates, on the other hand, necessitates the use of a pricing model that accounts for long-term impacts. Operator decisions on implementation, extension, and decommissioning, and fixed overhead costs will be included in such an energy market model. The latter argument is especially important: no power generation plant owner will construct more plants if energy profits only meet the marginal costs. The better-compared dismantling and investment costs of nuclear power production stations, for instance, are not adequately expressed in the merit order of ML design.

3.3. Forecast the consumer lifetime value

What is the consumer life value? The estimated amount of all possible sales (or profit margins) that a single consumer will produce for a company is known as consumer lifetime value (CLV or consumer LTV). The company's earnings will be maximized if reliable CLV forecasts are used as the basis for marketing decisions (or profits). Notably,

Table 9
ML mathematical techniques key benefits for the estimated cost of energy distribution.

Sr. #	Advantages area	Sub-class
1	Energy economics	Transmission and distribution capital savings Transmission and distribution operating and maintenance savings Energy efficiency Improved utilization of assets and energy infrastructure Energy theft reduction and detection at the distribution level Energy cost savings
2	Energy safety and security	Energy safety Energy security
3	Reliability of power systems	Control of power quality Control of power interruptions at different nodes of feeders, and disturbing networks
4	Climate concerns	Greenhouse gas emissions

[14] Merit order effect (MOE), https://www.next-kraftwerke.com/knowledge/what-does-merit-order-mean. Accessed: 03/12/2021.

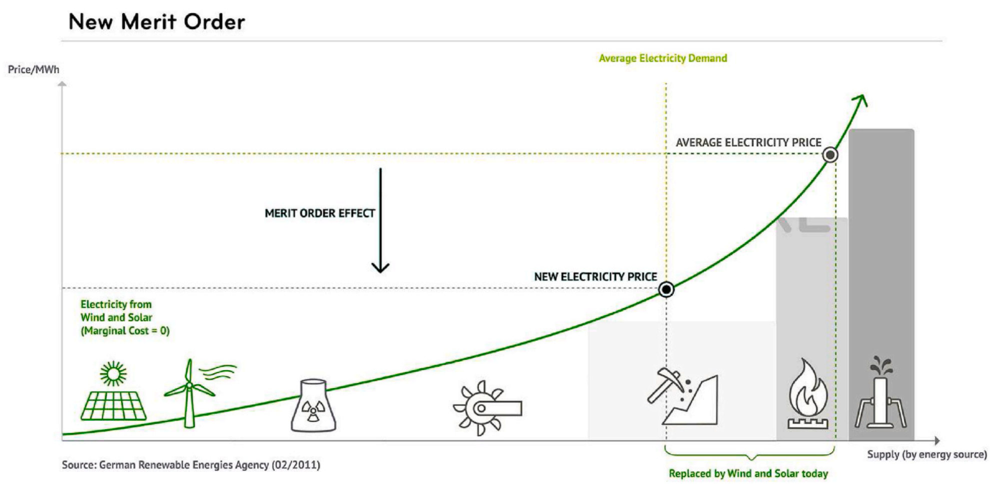


Fig. 16. Effect of the merit order on consumer energy demand.

estimating the correct consumer LTV forecasts is extremely difficult. Utility operators and suppliers must give greater attention to indicators like consumer lifetime value in an open services market [141,142]. This allows them to estimate how much each customer will pay throughout their agreement. ML does more than improve the customer lifetime value and prediction accuracy [143]. We can use ML algorithms like deep artificial neural networks to determine an individual consumer's total value by entering information like purchasing patterns, customer information, buying history, location, and payment behavior [144]. We may also go a step further with ML and recommend maximizing consumer satisfaction. This may include making extremely selective deals with related customers or using natural language processing (NLP) to boost support for unhappy customers on the verge of leaving.

3.4. Forecast the probability of winning consumers

To remain competitive in open markets, the power utility companies must have a complete record of future consumers [145]. On the other hand, ML will provide a more precise overview of consumer energy consumption. It also provides the data we use to enable data-driven marketing strategies [62]. This ensures power utilities able to tell when visitors or users visit the website [146]. The ML models used to compile the data the user carries with them – such as where they reside, what type of device they are using, search history, their browser history, and how often they have used the utility website – and create an accurate description of that individual as a customer. ML will then assess the probability of that individual being a consumer (recognized as scoring) and the best method for converting them into a consumer. This entails utilizing highly targeted advertisements and providing a highly customized interface – for example, displaying a screenshot of a family of four on the homepage when a family of four enters or opens the website.

3.5. Make high-targeted offers to consumers

Customers have a variety of service suppliers in competitive electricity markets [147]. Personalized or power utilities offerings are significant for attracting new consumers and retaining established ones, mainly because brand loyalty is not as high as it once was. ML helps determine the correct type of deal to make to a specific client at any particular moment by examining buying patterns and customer details. Suppose the data shows that a consumer is about to relocate. In that case, it may allow them to waive the connectivity charge at their new address. This kind of customized product keeps us ahead of the

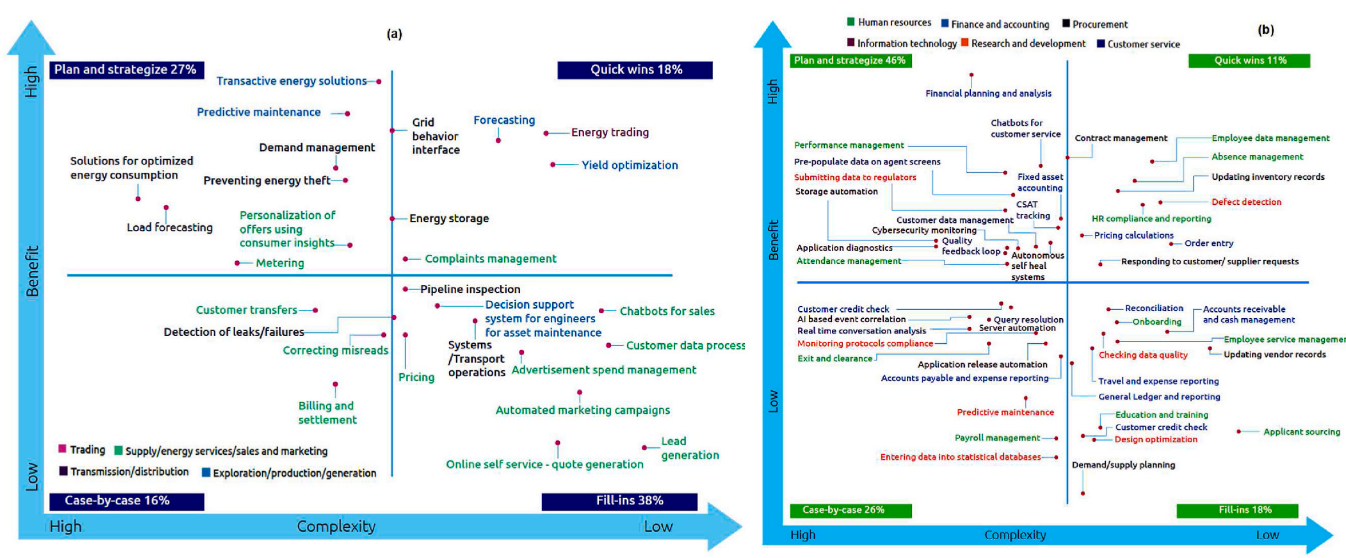


Fig. 17. Intelligent automation in utilities and energy distribution: (a) distribution of technical core-function cases by implementation complexity and advantages (b) distribution of support-function cases by implementation complexity and advantages.

competition and decreases churn probability with the customers.

Fig. 17 visualized the intelligent automation in utilities and energy distribution using ML models [51]. Fig. 17(a) reveals that more than one-third (38%) of energy and utility organizations make much effort to adopt low-benefit but straightforward cases. On the other hand, the emphasis is less than one out of five (18%) of quick wins. Just 11% of organizations rely on accelerated winning in support functions. Just over one out of four (26%) concentrate on use cases that are complicated to have and less convincing. It is noted that 46% of utility companies emphasize highly complex and beneficial applications which are used in real-time applications of power system infrastructure (see Fig. 17(b)) [51]. The percentage values indicate the implementation of use cases by utilities and energy organizations in each quadrant.

4. ML software for energy distribution

Bottom-up physics-based tools are used for urban building energy modeling and planning [148]. The urban modeling includes three basic classes, including 1) transport analysis and land use; 2) energy modeling and urban building; and 3) urban system energy modeling. Energy modeling and urban building are divided into major types, including i) top-down approaches and ii) bottom-up approaches. The famous tools for urban building energy planning includes: a) CitySim; b) SimStadt; c) umi; d) CityBES; e) OpenIDEAS; f) CEA; g) URBANopt; and h) TEASER. CitySim software was developed in C++ and Java language [149]. It aims to facilitate the design of sustainable urbanized areas. It can model the energy usage of a few to tens of thousands of buildings. SimStadt is an urban energy modeling tool intended to facilitate the energy transition preparation at the city level [150]. Energy ADE [151] and CityGML [152] are proposed in JavaScript and incorporate the 3D city authentication method. These tools are further used to explain the building's technical systems and fabrics. The urban modeling interface (UMI) was created to evaluate building energy usage at the community and city scales and sustainable public transport options, daylighting, outdoor convenience, and food security [153]. City Building Energy Saver (CityBES) tool was used to optimize and simulate the building energy efficiency/performance at a large-scale level [154]. OpenIDEAS is developed to use for a Modelica-based framework.

City energy analyst software is used to help manage energy consumption data [155]. Python developed energy Analysis and Simulation for Efficient Retrofit (TEASER) to integrate the USEM and UBEM [156]. A detailed analysis of this software used in different building energy modeling and analysis is given in Ref. [148]. Roberto et al. conducted a comprehensive study on energy assessment tools and methods [157]. Henrik et al. describe the EnergyPLAN that has been widely used in district cooling and heating systems and gas and electricity grids [158]. Fig. 18 (a & b) demonstrates the tools used for urban energy planning and management [159]. Energy planners are facilitated in selecting these tools based on skills, goals, and big data handling. As a result, energy planners determine the most appropriate method for a given application based on their expertise, goals, and data availability.

The power system software can be classified into three major classes: 1) power plants analysis software; 2) power engineering protection software; 3) and renewable energy controller software. The power plant analysis software is based on computations and mathematical algorithms. It includes ETAP, DIgSILENT, PSS/E, CYME, etc. Simulating protection mechanisms for power plants and power systems is another form of power engineering protection software. This software simulates the activation of different power system protections, which secure power lines, transformers, and other components. The renewable energy controller used different software, including DAC, ADC, 4-bit, 8-bit, 16-bit, and much more. The controller language is programmed with computer languages such as C++, C, Java, and others. Overall, software in the power system for energy distribution is used to model, simulate, and quantify power plants' configuration, transmission, electrical grids, lighting and grounding systems, and many other energy distribution structures. It is a program used to solve power engineering problems by converting them into mathematical expressions. Table 10 visualized the power system software used in different energy distribution systems.

Fig. 19 visualizes the power system simulator used to simulate the data and power flow by supplying energy to the miniaturized power grid.[15] Power system infrastructure includes substations, power plants, transmission loads, and lines, which are stimulated by ML models of miniature power equipment. The simulator is used to analyze the power grid fault conditions, real-time impact on the utility power grid, and redevelop a real-time electric power grid by flowing a small amount of electric current in the miniaturized power grid. These kinds of real-time simulations are useful for grid fault detection, the effect on the utility grid, and countermeasures with a high-level penetration of renewable energy. The data network that confirms international standards IEC61588 and IEC61850 using merging units and smart ML devices incorporated in each model makes it possible to calculate, control, and monitor the entire miniature grid just like an actual grid.

5. Challenges of machine learning for energy distribution

The preferences of consumers are moving to more automated customized services and applications. The flexibility of grid systems is rising behind-the-meter technology such as energy storage devices, hybrid cars, solar rooftops, and advanced metering technologies that enhances grid data availability and granularity. Simultaneously, competition from electricity distribution threatens firm models of utilities and accelerates a decrease in consumer loads. The most prominent challenges of ML models are listed below:

1) *Challenge of educating consumers:* Many businesses face the task of training their consumers in their advanced technology implementations. The same is true of ML engineers. The current capacity of ML is overestimated by entrepreneurs, programmers, and managers. They hope that the algorithms will easily understand and predict complex problems accurately.
2) *New technologies:* In reality, it is a relatively recent commercial application of ML, particularly in the field of deep learning. Users need a wide assortment of well-ordered and planned data to address questions. To answer questions, consumers need a broad range of well-ordered as well as scheduled data.
3) *Overfitting mechanism:* There are millions of variables in a standard ML network, perhaps hundreds of millions. Tens of thousands of records typically comprise a training data set. Although a network can recognize the training set and provide 100% precise and accurate responses, it can prove completely ineffective when new data are provided. The mechanism is referred to as overfits and is only one of the constraints of modern deep learning methods in ML.
4) *Talent deficit:* Although many individuals and the energy industry attract the ML market, few experts still have the opportunity to improve this technology. A good data scientist who has not enough experience in information engineering but who knows ML.
5) *ML development:* There are more layers in ML development. The engineers create software that learns to carry out the tasks expected to reach the company objectives. It is much more challenging to incorporate just one or two levels.
6) *Advanced energy technologies:* Advanced energy technology usually needs increasing advanced and complex materials for processes and systems. Complexity is a concern for conventional methods because of the large areas of the variable to investigate. Newly developed materials are generally needed for two decades,

[15] Power System Simulator, https://www.fujielectric.com/products/energy_ctrl_mng/b02.html, Accessed:03/12/2021.

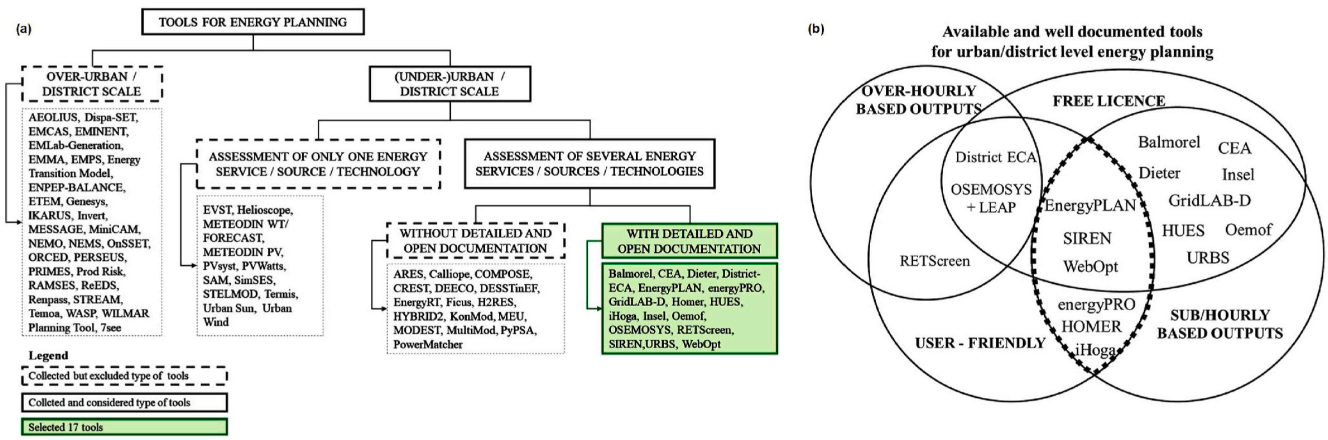


Fig. 18. Tools used for urban energy planning and management [159].

Table 10
Software used in power system protection.

Software name	Development history	Use of license (C=Commercial)	Latest available version	Developer	Purpose in energy distribution
NAP	1990	C	4.0.1	Innovation Energie Développement	Constrained and initial load flow, short circuit, stability, and contingency analysis calculation.
NEPLAN	1988	C	10.8.1.2	NEPLAN AG	Power system analysis, cloud computing, power system management, real-time integration, grid code, SCADA/GIS integrations, transmission and distribution networks, and assessment management.
SKM	1972	C	8.0.2.5	SKM Systems Analysis, Inc	HI-WAVE, TMS, CAPTOR, IEE Wiring, and IEC 60909 Fault.
CYME	1986	C	16.01	CYME International	Voltage stability analysis and COM module.
PSSE	1976	C	–	Siemens	Over timescales of a few seconds to tens of seconds, as well as in steady-state situations.
EMTP	1982	C	4.1	EDF & RTE & Hydro-Québec	Electromagnetic transients' programs, insulations issues.
ERACS	1990	C	3.9.10	RINA Consulting Ltd	Modeling tool balanced 3-phase power systems stability analysis which includes load flow of power systems, short/fault-circuit, protection coordination, harmonics and G5/4, arc flash calculation modules, and transient stability.
PSCAD	1986	C	4.003	Manitoba HVDC Research Centre	Short circuit analysis, power system optimization, cable pulling, power flow study, DC and AC arc flash measurement, contingency, and voltage stability analysis, estimation of transmission line parameters, power system control, and optimization, proactive device coordination.
XGSLab	2004	C	7.01	SINT Ingegneria	GSA, XGSA FD, GSA FD, XGSA TD.
DIgSILENT	1985	C	2018	Dr. Martin Schmieg	Grid code, power factor 2018, and station ware 2018.
ETAP	1986	C	19.0.1	Operation Technology, Inc.	Power plant controller, power system management, microgrid controller, SCADA, geospatial modeling, EMS, ADMS, transmission, and distribution planning.

from discovery to implementation. Further, timing error prohibits new materials from being incorporated into products efficiently, compromising efficiency and competition.

7) *Technology developments:* Achievable performance goals, established ambitiously, and a comprehensive research and development portfolio to achieve these goals is also challenging.
8) *Energy storage devices and materials:* The growing complexities of energy storage devices and energy storage systems and the vast volume of background data present considerable challenges for standard techniques and algorithms. New cutting-edge technologies overcome the challenges of the conventional methods for increased precision, reliability, and optimization.
9) *Technology challenges:* Firstly, energy storage development includes innovation and breakthrough, long-term storage, a high level of security for electrochemical storage, and low cost. Furthermore, high efficiency and physical storage technology need low cost. The second point is that research concentrates on simulating energy storage and optimizing the process in various energy systems, theoretically supporting the utilization of energy storage technologies and creating experimental setups and detailed evaluations for industrializing and marketing energy storage. Complete and robust cohesion, appropriate definition, accountability, openness, and energy storage requirements should also be established, offering good support to research and growth, storage applications, and energy generation, and fostering the energy storage development technologies and related industries that are yet to be challenging.
10) *Economic Challenges:* Today, some countries' energy storage sector faces challenges such as lack of support for the legislation, high costs, uncertain value, unhealthy business mechanism, and other issues [160]. In the future, it would be important to take into consideration two aspects: firstly, the proposal of alternatives to the energy storage scheme involving power generators, electrical companies, researchers, economic and social organizations and, secondly, the promotion of an appropriate business competition structure and subsidy strategy for new ML innovative technologies. Based on Woori's prediction, the world energy storage cost will rise 26% per year in the future, and the total

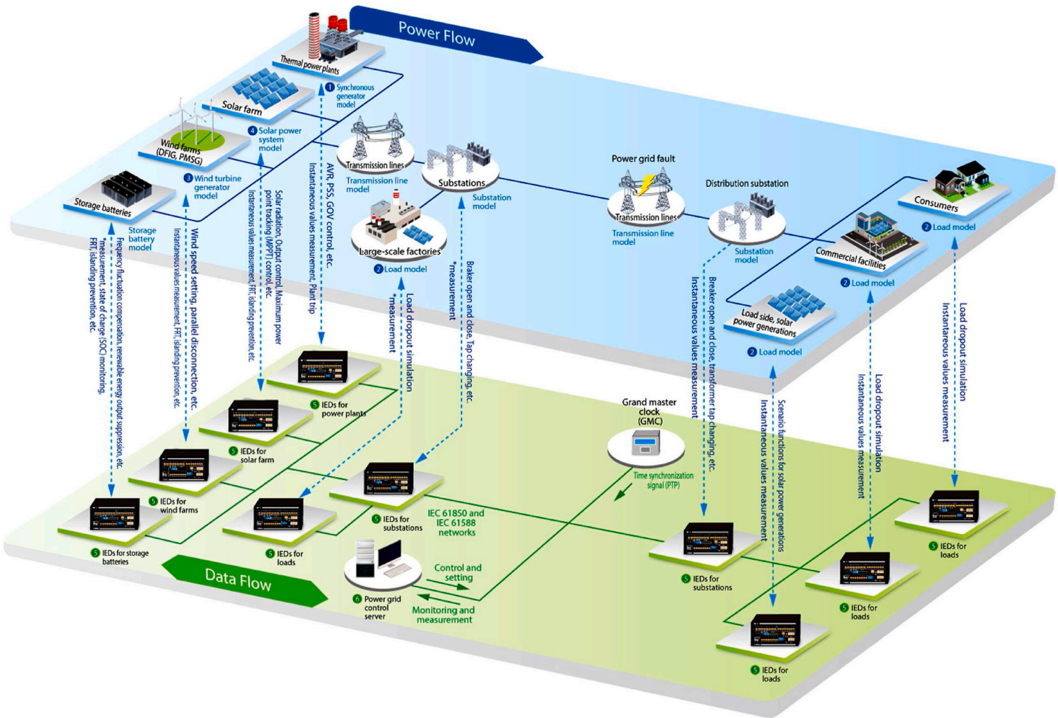


Fig. 19. Real-time power system simulator.

energy storage market value in 2020 is noted up to $16 billion [160]. While energy storage has several market dynamics, the key hurdles remain high costs, inadequate subsidy policies, an indeterminate pricing structure, and the business model.

11) *Energy-efficiency perspective:* ML strategies' promise is great in addressing future green 5G and 6G energy management challenges. ML's methodologies such as federated learning, deep learning, and optimization may be discussed for the design and optimization of the network orchestration and architecture cost-effectively. ML learned network complexity for designing 5G and 6G air interfaces by acquiring dynamic network topology and varying traffic patterns. ML would be more far-reaching and critical as saving electricity employing diversified 5G and 6G technologies, such as smart grid, smart transmission and distribution networks, smart cities, and factory automation. On the other end, coordination and computing are generally needed in ML techniques. This could present a major obstacle for the design and deployment of energy-efficient ML algorithms as well as of future 5G and 6G networks in smart energy systems [161].
12) *Demand response:* First, demand response agents need to operate in a partially observable atmosphere, i.e., agents are unable to fully understand the units that were used for demand response, other environments, and agents [162].
13) *Intermittent nature of renewable energy:* Power utilities have challenges and prospects for innovation and new technology in the energy sector. The intermittent nature of renewables requires the pace of decision-making to leverage renewables that contradict the conventional prediction models. In the meantime, the deployment of intelligent meters provides a thin and large energy usage. Efficient processing and use of this information may provide new load forecasting prospects.
14) *COVID-19 Impact on energy demand and requirement:* The coronavirus pandemic provides an understanding of the energy use and thus highlights vulnerabilities within utilities' environments as one of the new challenges that utilities face. Poor predictions or a general lack of forecasts negatively impact public infrastructure, leading to waste of energy, higher operating costs, potential outages of electricity, and financial loss. The error in the ML forecasting model lacks the overall system efficacy and performance. Increased accuracy and an understanding of these improvements to grid loads prevent cost-effectiveness errors and make smarter grid management decisions.
15) *Solar Energy:* Solar power systems and associated technologies have become renewable energy sources. Their deployment at the international level is extensively recognized. Due to the relative cost of installation,

poor conversion efficiency and battery capacity problems, solar power is still not a widespread source of electricity compared with conventional power sources. Despite the difficulties in deploying ML techniques, there is much groundbreaking research on emerging practices such as AI, deep learning, big data, IoT, and new approaches to enhance the performance of solar power transformation efficiency in the field of energy are needed. Many companies are working to overcome the above-said challenges with the use of ML in energy distribution. Table 11 demonstrates the consortiums, companies, and foundations working on ML in energy distribution. According to our best knowledge, just a few companies are listed in Table 11.

6. ML opportunities towards a smart and sustainable future

The combination of resilient, low latency connectivity, high-speed, and technologies such as the ML will transform towards sustainable smart energy industry distribution. With the environmental effects of power generation and consumption growing, utilities aggressively pursue novel alternatives to reduce their environmental effects. ML is currently being integrated by leaders of both the traditional and renewable energy markets in many ways to increase alternative energy use while improving the accuracy of renewable energy forecasts, energy conservation, and usability. Among them, key points of ML towards a smart and sustainable future are listed below:

1) *Energy investors:* In the context of an ever more volatile economy, energy investors looking towards resilience are re-developing ML

Table 11
Consortiums, companies, and foundations working on ML in energy distribution [163].

Sr. #	Name of the company	Company providing services	Objective of services
1	BeeBryte (Singapore, and France)	Demand-side management and energy demand forecasting	BeeBryte is intended to minimize the electricity using ML algorithms for energy charges and automatic controls on heat-cooling devices (for example, HVAC), generators, charging points for electric vehicles, or batteries.
2	IBM Watson (The United States of America)	Power grid reliability, and stability	IBM uses analytics to help make power-decision decisions. IBM balances supply and demand for sustainable, stable, and efficient electricity services from traditional and renewable sources. It monitors and manages grids globally.
3	DeepMind, Google (The United States of America)	Demand-side management and energy demand forecasting	DeepMind designs programs to solve the complex nature of power system problems. In an attempt to reduce power consumption, DeepMind tested its ML techniques on Google's data centers.
4	Tomorrow (Denmark)	Demand-side management and energy demand forecasting	The ML models were developed by Tomorrow that systematically derives insights from different types of data on CO2 emissions. Such observations are then used by other resources, including the ElectricityMap, which shows the CO2 emissions from the development, export, and import of energy in various regions worldwide.
5	Verv (United Kingdom)	Demand-side management and energy demand forecasting	To learn about home energy products and their behaviors, Verv energy assistant aims to reduce the household's energy costs by using ML.
6	DCbrain (France)	Grid reliability and stability	DCbrain allows flow and consumption optimization, power system network and anomaly detection and prevention, and network evolution simulation.
7	EUPHEMIA, N-SIDE (Europe)	Optimized energy market operation	EUPHEMIA is a method of coupled energy integrated into European energy markets for determining spot volumes and prices in the European energy markets
8	Fraunhofer (Germany)	Power grid reliability and stability	The Fraunhofer Institute has developed an ML algorithm capable of logging and containing up to 4.3 million data sets every day, analyzing these data to reliably forecast network operators, identifying and acting on network abnormalities within 20–50 ms.
9	SmartNet (European Union)	Power grid reliability and stability	SmartNet offers resources to enhance cooperation between distribution system operators and transmission system operators by sharing information on the control and the collection of auxiliary services from district actors.
10	McKinsey, Utilityx, (The United States of America)	Predictive maintenance control	Utilityx supports infrastructure management in predictive engineering to optimize efficiency. Advanced analytics was used to turn network data into a health-driven and essential asset-based approach.
11	Kunumi and PSR (Brazil)	Market operations and advanced energy demand forecasting	Kunumi and PSR combine ML and modern theoretical approaches, including energy planning, operation, and trade, to maximize and anticipate uncertainty in energy systems.
12	Infosys (India)	Demand-side management and energy demand forecasting	By applying ML to data generated by the smart meters, advanced sensors, and smart devices behind meters, Infosys supports stakeholders from the energy sector.
13	Grid Edge (United Kingdom)	Grid reliability and stability	The company grid edge renders the cloud-based ML models which allow to optimize, forecast, and control the supply energy demand.
14	EWeLiNE (Germany)	Renewable energy production and forecasting	EWeLiNE developed the ML models, which are used for renewable energy demand forecasting. EWeLiNE collects real-time wind turbine and solar data from across Germany and incorporates it into a calculating method for renewable energy generation for the corresponding 48 h.

power to direct data-driven decisions. It also reduces the electricity consumption in a building or manufacturing facility with the same technology. ML is expected to be applied in several diverse fields of energy efficiency and the development of clean energy as well as sustainable future [37]. The key studies or investigations should aim to upgrade new energy technologies for energy conservation accompanied by renewables' increasing use. Recent technical advancements in ML technology have allowed applying ML methods in the energy and environmental sectors to guarantee sustainable growth [164].

2) *Smart grid and distribution generation* [165]*:* As an important part of energy distribution, global warming, and climate change impose a significant shift in the use of resources and more sustainable use of renewable resources: development, supply, and consumption. In the energy transformation era, smart grids should improve power networks and systems. The energy market has changed inexorably with the application of integrated grids, microgrids, and technological innovations. The most interesting factor is that smart grids maintain safety in supply and allow customers to engage as prosumers in the

energy market. Smart grids use automated and advanced technologies to control and track power transmission from all generating sources to meet end-user demand rapidly, quickly, and efficiently. Intelligent smart grids improve flexibility, device durability, and stability and minimize the climate's disturbances, expense, and effects. Some emerging technologies, including distributed generation and microgrids, deliver electricity locally, create broader and more stable networks, and lower line overload. Electricity storage complements renew energy while local energy supplies contribute to the micro-grids reduction of all blackouts. The microgrid is fitted with local storage and power supply infrastructure based on independent distributor power supplies. To allow owners to produce their electricity, microgrids often reduce their energy dependence by contributing to cost reduction and avoiding peak consumption costs. The microgrid will generate income if it produces an excess of electricity allocated or supplied to the energy utilities. Whatever is said above is possible only due to advanced ML technologies.

3) *ML reshaping the energy industry* [92]*:* Renewable energy will be supplied by 2050 up to 70% of the world's total energy demand [166]. The method we consume electricity is simultaneously evolving. Large consumption of electric cars is projected in the coming years. Digitalization enables these key patterns of decarbonization and decentralization. This reorganization of the energy distribution is made possible by modern ML technologies.
4) *Data-driven strategies* [167]*:* Data-driven strategies with the deployment of ML in the utilities and power industry renders the platform for management discussion around the implementation of data-driven management and data-science to deliver energy strategy in long-term perceptive, and consequently, competitive advantage for a sustainable future.
5) *Decision making* [136]*:* ML can enhance AI system efficiency for energy distribution decision-making. Reinforcement learning can be used in any situation requiring decision-making in an inconsistently changing environment. Machine control, robotics, and tuning (i.e., heating, ventilation, and air conditioning), optimization of the supply chain (i.e., ML demonstrates how much of a given object to buy from when and where), and crop yield optimization are all examples. Reinforcement learning (RL), a sub-class of ML, nearly often generally requires instruction in a simulated environment. *How will RL be used to meet the Sustainable Development Goals?* By training robots, drones, and smart devices how to make better choices in diverse use cases, RL will help us meet any or all of the Sustainable Development Goals (SDGs), for illustrate, (Sustainable Development Goal 3 for well-being and health), (Sustainable Development Goal 9 for infrastructure and industry innovation), (Sustainable Development Goal 12 for production and responsible consumption), etc.

7. Recent progress on properties and discovery of machine learning in the energy distribution

It is widely expected that ML can change almost every part of our society. The most commonly used subset of AI today is ML techniques. Deep learning models or a subset of ML utilizes multi-level neural networks in ML to learn from the massive amount of data stored from different energy sources, devices, and power system infrastructure. The key point on the recent progress of ML in energy distribution is listed point-by-point below:

1) *New energy materials* [7]*:* Recently, the screening of new materials and their simulation of quantitative relations of structural behavior have become a hot and trendy subject in energy materials because of the many opportunities and challenges, including a low probability of performance, high computational costs, and time consumption connected with classical techniques of energy material development [7]. Currency algorithms for regression, classification, clustering or dimensional reduction of large collections of particularly high-dimensional input data are now being used effectively. Therefore, recent advances in ML have raised the probability that data-orientated materials research would transform breakthroughs to provide innovative energy material development strategies. In addition, recent developments in data-driven materials science also show that ML technologies' use would make it much simpler and accelerate the discovery and delivery of state-of-the-art energy materials [168,169]. ML applications is used in different energy materials, for illustrate, alkaline ion battery materials (e.g., electrolytes, electrodes, etc.) [170], catalytic materials [171,172], photovoltaic materials (e.g., property screening and forecasting, solar conversion efficiency, organic PV, etc. (See Fig. 20)) [38,173,174], carbon dioxide capture materials [175,176], and catalytic materials [177].
2) *Energy storage technologies:* In new energy applications such as integrating renewable energies and peak load shaving technologies, energy storage technologies are well recognized. Fig. 21 visualized the operational power range versus the time of many new energy storage technologies. It demonstrates the suitability of many utility applications. For renewable energy integration and utility, life cycle, output power, and energy storage capacity are the key criteria of material capacity and storage performance. In particular, the use of ML and the proliferation of storage would allow renewable energy to be incorporated and dispatched to promote smarter grids that depend less as well as inefficient power generation plants.
3) *The United States Energy Department:* The United States National Laboratories and Energy Department established an environment of research to achieve their scientific, technology, and safety missions using ML techniques. This environment expands and enhances ML programs, software, algorithms, processes, and applications. DOE and its labs work together to promote technology advancement that ensures stability in the United States and competition with other federal departments, the industry, and academia [179].

The United States Energy Department proposed different strategies towards the use of ML in energy distribution. ***Strategy-1:*** initiate long-term investments in ML area for energy distribution; ***Strategy-2:*** the development of effective techniques for human-ML collaboration; ***Strategy-3:*** address as well as understand the ethical, legal, and societal repercussions of ML; ***Strategy-4:*** ensure the security and safety of ML systems; ***Strategy-5:*** proposed shared environments and shared public datasets for ML testing and training; ***Strategy-6:*** evaluate and measure the ML techniques through benchmarks and standards; ***Strategy-7:*** better understanding for research and development of ML techniques; and ***Strategy-8:*** expansion of the private and public partnership to accelerate the advances in ML techniques.

4) *The use of ML in Energy Science:* The science office has allocated 13 million US dollars to the development of ML as a platform for forecasting and scientific investigation. Developing applications and algorithms for particular scientific programs would cost about $11 million.[16] The remaining funds will be used to boost the predictability of ML models. The electricity office reported federal support for many ML-related initiatives totaling $7 million. These initiatives are intended to help in the application and development of faster grid simulation and analytics in the future.[12] The "Advanced Research Projects Agency-Energy" has launched a new initiative called design intelligence for energy

[16] Department of Energy, Department of Energy Announces $20 Million for Artificial Intelligence Research, https://www.energy.gov/articles/department-energy-announces-20-million-artificial-intelligence-research, Accessed: 03/12/2021.

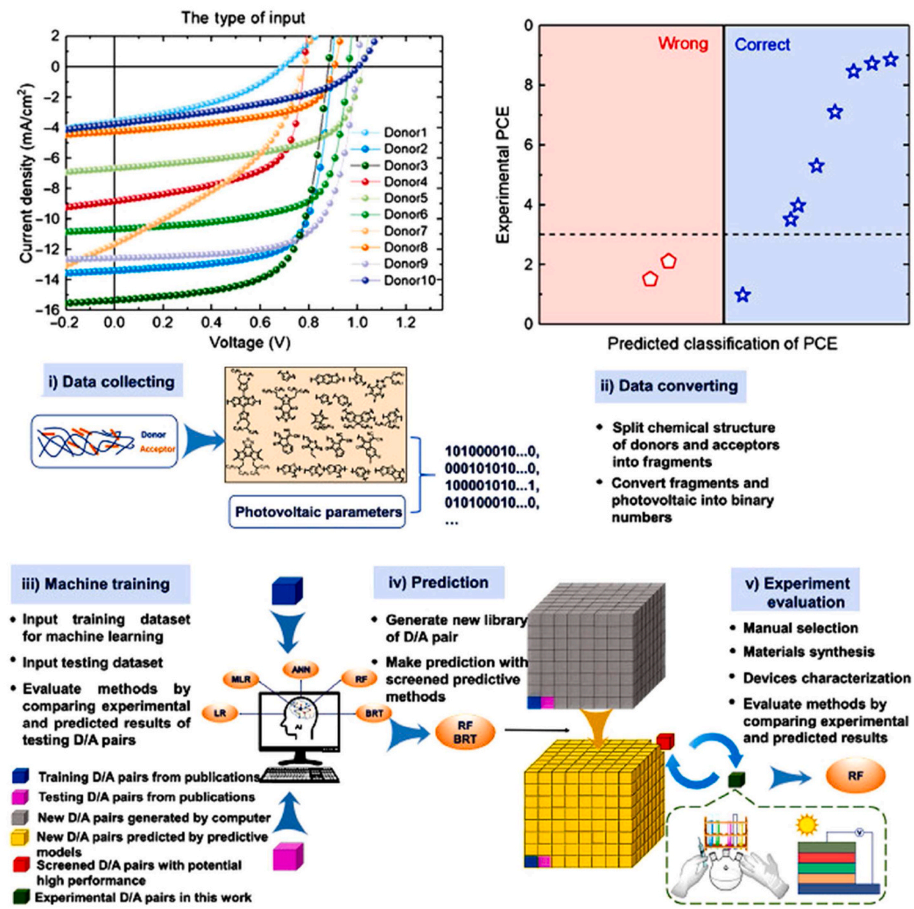


Fig. 20. The processes of ML emerging technologies for advanced organic solar cell stages of development [7].

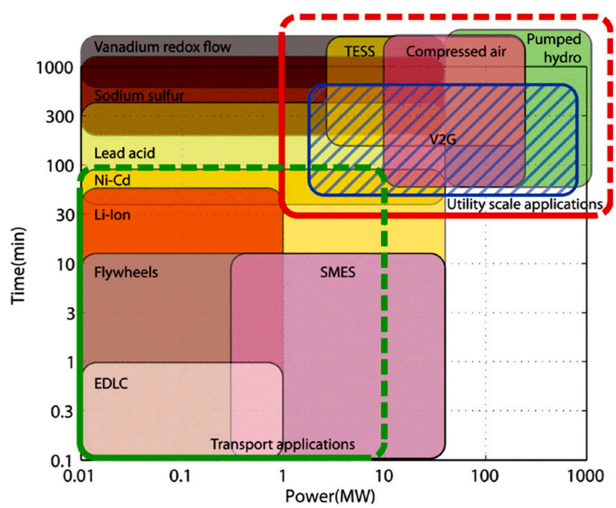


Fig. 21. Applications of energy storage technologies for energy utilities [178].

consumption reduction and advanced technology upgrades, which would grant up to $20 million to projects.[17] ML can be implemented in energy systems by project teams and instigating energy technology development.

5) *Renewable Energy and Grid Modernization* [180]*:* The United States DOE and its national laboratories are taking advantage of recent developments in ML to further use their vast computing capabilities and technological workforce in favor of grid automation and energy conservation, and sustainable energy utilization.
6) *National Security and Cybersecurity* [37]*:* ML technologies built by the United States DOE and its National Laboratories prevent malicious cyber-attacks on critical energy infrastructure, and reduce the grid disruptions and shutdowns. Preventing power failures has major social and economic implications around the world. The ML technologies facilitate the modernization of the power grid using autonomous energy systems. The long-term objective is fully achieved through key objectives: 1) secure; 2) flexible; 3) scalable; 4) robust; 5) resilient; 6) affordable; 7) reliable; and 8) real-time interoperability.
7) *Geothermal:* ML techniques that are successfully implemented could aid in the exploration of geothermal wells, improve drilling accuracy, and lower costs.[18]
8) *Biomass:* The Idaho National Laboratory scientists utilize the ML models for data analysis on bio-refinery processing to direct operational design changes that optimize performance while mitigating device damage.[19]

[17] Design Intelligence Fostering Formidable Energy Reduction and Enabling Novel Totally Impactful Advanced Technology Enhancements, https://arpa-e.energy.gov/technologies/programs/differentiate, Accessed: 03/12/2021.

[18] Office of Energy Efficiency & Renewable Energy, Energy Department Awards $5.5 Million to Apply Machine Learning to Geothermal Exploration, https://www.energy.gov/eere/articles/energy-department-awards-55-million-apply-machine-learning-geothermal-exploration. Accessed: 03/12/2021.

[19] Artificial intelligence helps turn biomass into energy, https://inl.gov/article/systems-engineering/. Accessed: 03/12/2021.

9) *ML Supercomputing Capabilities:* The US DOE has success in the national laboratories' world-class supercomputing capabilities. The labs are also the home of four of the world's ten fastest supercomputers – establishing the US as the worldwide leader in the area of high-tech ML computing. The most powerful supercomputers include: 1) Cascade (e.g., 3 Petaflops); 2) Cori (e.g., 28 Petaflops); 3) Mira (e.g., 10 Petaflops); 4) Joule 2.0 (e.g., 6 Petaflops); 5) Sequoia (e.g., 20 Petaflops); 6) Lassen (e.g., 23 Petaflops); 7) Theta (e.g., 7 Petaflops); 8) BeBop (e.g., 2 Petaflops); 9) Summit (e.g., 200 petaflops); and 10) Trinity (e.g., 41 petaflops). Exascale computing is the next generation of computing systems. The Exascale computer can perform quintillion computations per second, demonstrating the unprecedented breakthroughs in ML. More detailed analysis and recent progress concerning ML techniques' supercomputing capabilities are given in Ref. [179].
10) *Technology Advancement:* To simulate the electrical system and the complexity of different issues much faster than possible with current commonly used techniques, Argonne scientists are working on optimization methods that use ML and AI. The key emphasis is to speed up the electric system planning and daily load flow analysis and calculations.
11) *Demand Response* [181]*:* The demand response management and infrastructure are intended to track and use the real-time energy consumption information to deliver the energy prices to thousands of consumers in real-time via the utility power grid. Customers may respond to these rates and enable them to change their energy consumption in terms of gird conditions. Continued development can boost stability, cost savings, and sustainability by helping end-user applications consider how they need power grid changes. Information and such resilience would encourage the potential penetration of renewable energy into the power grid.
12) *High Fidelity:* Argonne National Laboratory increases high-fidelity simulation with ML, such that the optimization of the design process is accelerated significantly while sustaining the data durability. In this scenario, a task or job that takes hours using high-fidelity alone takes a milliseconds optimization process augmented by ML.
13) *Smart Energy Materials Manufacturing* [38]*:* Advanced processing techniques require multi-component structures and intricate thermo-chemical processes resulting in a high data volume, often at a high pace. Optimally, operators can get immediate data/feedback on the characteristics of manufacturing of energy materials and process variables in real-time to find new processes and phenomena faster and adapt to increase effectiveness and efficiency. However, existing methods offer "postmortem" data/feedback long after the manufacturing process is completed. ML methods are used to optimize the analyze the manufacturing process.
14) *Solar Energy breakthroughs with ML* [58]*:* There are five directions where ML takes advantage of advances in the field of solar technology. These areas of research include 1) smart infrastructure design for solar energy forms and systems (e.g., optimal sizing, weather conditions, accurate geographical placement, etc.); 2) smart maintenance of solar power plants (e.g., advanced sensors, failure prediction, anomaly detection, and real-time power system data) [182]; 3) solar energy generation forecasting (e.g., historical satellite data, environmental and real-time forecast information can be integrated into and evaluated using optimization analytics to support hardware maintenance decisions) [183]; 4) optimization process of power transmission and distribution networks (e.g., integration of energy consumption data [184], proactive monitoring of distributed grid, etc.) [185]; and 5) well understanding of solar energy markets (e.g., consumption patterns, price fluctuations, production metrics, and historical prices, etc.) [186].

8. Conclusion

The supply of continuous energy to consumers has many challenges to address for power grids. Energy usage, weather patterns, even wildcard incidents, internal failure, interferences, and lightning strokes from wild animals can impact power delivery in energy distribution networks. The ML is increasingly being applied to help forecast potential brownout conditions. This study covers the detailed review analysis for data-driven probabilistic ML and its application for energy distribution. ML models have been applied in a variety of applications for energy distribution. Power utilities and grid operators use ML to classify grid failure conditions at generation, transmission, and the distribution level. Smart meters and sensors such as phase measurement units render energy consumption information in real-time. When we combine ML tools with simulated and historical data, ML is used to help and mitigate grid failure conditions, using different models such as demand-response optimization and grid balancing. Furthermore, recent advances in data-driven systems engineering show that ML techniques can be used to facilitate and improve the deployment and development of advanced energy materials and electric infrastructure at the energy distribution level.

ML models are used to figure out the fault conditions more precisely, supporting consumers to mitigate service interruption. For example, discrete wavelet transforms with a SVM are used to locate the fault condition at energy distribution networks and systems. The grid balance - ensuring that electricity supplies meet energy demand - is one of the main tasks for energy utilities and grid operators. However, weather-dependent renewable sources make energy forecasting even harder. The most accurate tool for forecasting the demand for renewable energy generation is ML in energy distribution. ML leads to many weather variables (e.g., wind speed, wind direction, solar radiation, etc.) and makes perfect energy forecasts by applying specialized methods such as short-term neural networks. This saves operators money and conserves power plant energy.

More and more cyberattacks are being used to target important power system infrastructure, such as hijacking the power system control and demanding money. Utilities and grid operators use ML to distinguish between a malfunction (for illustrating, short circuit faults) or a disruption in the power grid and a smart cyber-attack (such as false data injection to the power system). Symbolic dynamic filtering theory is used for feature extraction to explore causal interactions between the subsystems without overburdening computer systems. When tested at power distribution systems, 99% of cyber-attacks were correctly identified with a true positive rate of 98% and a demonstrably false rate of less than 2%.

The use of ML models has been well recognized in energy market information management (e.g., financial assurance, market mitigation and monitoring), market operations (e.g., economic dispatch, unit commitment, security constrained, energy reserve calculation and requirements, system black start capability, voltage control and reactive supply), power system modeling (e.g., data acquisition, supervisory control, and power system modeling), outage and forecast (e.g., climate forecast, energy demand forecast, and outage scheduling), transmission scheduling (e.g., tagging, external scheduling, transmission capacity measurement, and energy distribution network reservation), grid operations (e.g., dispatch management, state estimation, dynamic limit estimation, automatic energy production control, resource performance monitoring, voltage stability analysis, resource capacity analysis, flow gate calculation and power flow, and contingency analysis), grid settlements (e.g., internal scheduling, billing, energy tariffs, metering, reliability, and ancillary services), utility organization (e.g., customer data management), energy distribution control (e.g., load monitoring, alarm processing, and logging), energy security (e.g., policy management, energy authentication and authorization), energy demand and consumption data management (e.g., data visualization, business intelligence, data archive and historian).

Recent progress in science and engineering data-driven materials has shown that ML technologies' implementation will make the design, research, deployment, and development of advanced energy materials much easier. This study explores that is the most widely used ML techniques in energy distribution development and research include supervised learning (e.g., support vector regression, SVM, RVM, decision tree/RF, and regression models), unsupervised learning, and deep learning (for illustrating, EML, artificial neural networks, RNN, CNN-based methods). Several modern ML techniques combine multiple methods and algorithms (i.e., integrating RNN and CNN, combining DNN and RNN) to achieve the highest level of training, test, and validation learning performance.

The key breakthrough of successful ML application in energy distribution and remaining challenges is pointed out like standardization and improvement of energy infrastructure, model visualization and automatic closed-loop optimization, experimental exploration, supporting policies, and interdisciplinary communications. The energy industry moves from a reliable, low-risk environment, regulation-driven to a modern technology-driven, more sophisticated, as well as uncertain marketplace with the use of ML models. Changes, as well as emerging technology and business models used in energy distribution, would be essential. Unrecognizing this could threaten the viability of power companies. Our study findings show that ML can play an important role in navigating and managing this emerging energy distribution system. It will probably be one of the key enablers. However, the utilities seem reluctant to adopt this available resource (e.g., ML techniques) due to unknown uncertainties. We believe this is a mistake. As such, this study's recommendations illustrate how power companies can take advantage of the benefits of ML and keep ahead of the ML curve relatively quickly. But to do this, they have to become smart and act quickly.

Credit author statement

Tanveer Ahmad: Conceptualization, Methodology, Data curation, Writing – original draft, Visualization, Validation, Writing- Reviewing and Editing. Rafal Madonski: Methodology; Writing- Reviewing and Editing. Dongdong Zhang: Investigation; Methodology; Writing – original draft. Chao Huang: Conceptualization; Funding acquisition. Asad Mujeeb: Data curation; Methodology.

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.

Acknowledgments

This work was supported in part by the National Key Research and Development Program of China (Grant No. 2019YFE0118000), the Key Laboratory of Special Machine and High Voltage Apparatus (Shenyang University of Technology), Ministry of Education, Grant KFKT202006, and Guangxi Young and Middle-aged Scientific Research Basic Ability Promotion Project, 2020KY01009, National Natural Science Foundation of China under Grant 62002016, in part by Beijing Natural Science Foundation under Grant 9204028, in part by Guangdong Basic and Applied Research Foundation under Grant 2019A1515111165. The authors are grateful for the sponsorship of the UM-Macau Postdoc Talent Program under the cover of the Government of Macau, SAR.

References

[1] McCollum DL. Machine learning for energy projections. Nat Energy 2021:1–2. https://doi.org/10.1038/s41560-021-00779-9. News & Vie.
[2] Lu T, Sherman P, Chen X, Chen S, Lu X, McElroy M. India's potential for integrating solar and on- and offshore wind power into its energy system. Nat Commun 2020;11:1–10. https://doi.org/10.1038/s41467-020-18318-7.
[3] Zhang Weiping, Maleki * Akbar, Rosen Marc A. A heuristic-based approach for optimizing a small independent solar and wind hybrid power scheme incorporating load forecasting. J Clean Prod 2019;241:18.
[4] Schiermeier Q. And now for the energy forecast. Nature 2016;535:212–3.
[5] Omer AM. Focus on low carbon technologies: the positive solution. Renew Sustain Energy Rev 2008;12:2331–57. https://doi.org/10.1016/j.rser.2007.04.015.
[6] David TW, Anizelli H, Jacobsson TJ, Gray C, Teahan W, Kettle J. Enhancing the stability of organic photovoltaics through machine learning. Nano Energy 2020;78:105342. https://doi.org/10.1016/j.nanoen.2020.105342.
[7] Liu Y, Esan OC, Pan Z, An L. Machine learning for advanced energy materials. Energy AI 2021;3:100049. https://doi.org/10.1016/j.egyai.2021.100049.
[8] IRENA. Renewable energy highlights. 2020. p. 1–2. https://doi.org/https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2020/Jul/Renewable_energy_highlights_July_2020.pdf?la=en&hash=75B114DB7A55F4260F41F64C4DFF793DB2044306.
[9] Mai T, Hand MM, Baldwin SF, Wiser RH, Brinkman GL, Denholm P, et al. Renewable electricity futures for the United States. IEEE Trans Sustain Energy 2014;5:372–8. https://doi.org/10.1109/TSTE.2013.2290472.
[10] DSIRE. Database of state incentives for renewables & efficiency. NC Clean Energy Technol Cent NC State Univ; 2021. . [Accessed 28 February 2021].
[11] (SEPA) SEPA. Utility carbon reduction tracker. Sepa 2020 [https://doi.org/https://sepapower.org/utility-carbon-reduction-tracker/].
[12] Kuhn T. Delivering America's energy future. Elec Perspect 2018;43:6–7 [https://doi.org/https://www.eei.org/issuesandpolicy/finance/wsb/Documents/EEI_WSB_Remarks.pdf].
[13] IRENA. REmap 2030 - a renewable energy roadmap./Publications/2014/Jun/REmap-2030-Full-Report 2014:188. https://doi.org/https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2014/IRENA_REmap_summary_findings_2014.pdf.
[14] Mutezo G, Mulopo J. A review of Africa's transition from fossil fuels to renewable energy using circular economy principles. Renew Sustain Energy Rev 2021;137:110609. https://doi.org/10.1016/j.rser.2020.110609.
[15] Antonopoulos I, Robu V, Couraud B, Kirli D, Norbu S, Kiprakis A, et al. Artificial intelligence and machine learning approaches to energy demand-side response: a systematic review. Renew Sustain Energy Rev 2020;130:109899. https://doi.org/10.1016/j.rser.2020.109899.
[16] IRENA. Innovation landscape for a renewable-powered future: solutions to integrate variable renewables. Abu Dhabi: Int Renew Energy Agency; 2019. p. 1–164.
[17] IRENA. Global energy transformation. Int Renew Energy Agency 2018:1–76.
[18] Murphy KP. Probabilistic machine learning: an introduction (adaptive computation and machine learning). MIT Press; 2020. p. 1–932.
[19] Murphy KP. Probabilistic machine learning: an introduction (adaptive computation and machine learning). MIT Press; 2020. p. 1–932. Massachuse.
[20] Cheng L, Yu T. A new generation of AI: a review and perspective on machine learning technologies applied to smart energy and electric power systems. Int J Energy Res 2019;43:1928–73. https://doi.org/10.1002/er.4333.
[21] Donti PL, Kolter JZ. Machine learning for sustainable energy systems. Annu Rev Environ Resour Oct. 2021;46(1):719–47. https://doi.org/10.1146/annurev-environ-020220-061831.
[22] Sharmila P, Baskaran J, Nayanatara C, Maheswari R. A hybrid technique of machine learning and data analytics for optimized distribution of renewable energy resources targeting smart energy management. Procedia Comput Sci 2019;165:278–84. https://doi.org/10.1016/j.procs.2020.01.076.
[23] Perera ATD, Kamalaruban P. Applications of reinforcement learning in energy systems. Renew Sustain Energy Rev Mar. 2021;137:110618. https://doi.org/10.1016/j.rser.2020.110618.
[24] Westermann P, Evins R. Using Bayesian deep learning approaches for uncertainty-aware building energy surrogate models. arXiv:2010.03029 [cs, stat]. Oct. 2020 [Online]. Available: http://arxiv.org/abs/2010.03029. [Accessed 21 November 2021].
[25] Sun M, Zhang T, Wang Y, Strbac G, Kang C. Using bayesian deep learning to capture uncertainty for residential net load forecasting. IEEE Trans Power Syst Jan. 2020;35(1):188–201. https://doi.org/10.1109/TPWRS.2019.2924294.
[26] Pfingsten T. Fully bayesian inference and model selection for Gaussian process regression. 2019. https://doi.org/10.13140/RG.2.2.21150.38727.
[27] Munkhammar J, Widen J, Grahn P, Ryden J. A Bernoulli distribution model for plug-in electric vehicle charging based on time-use data for driving patterns. In: 2014 IEEE international electric vehicle conference (IEVC), florence; Dec. 2014. p. 1–7. https://doi.org/10.1109/IEVC.2014.7056224.
[28] Coulier A. Multiscale modeling in systems biology: methods and perspectives. Acta Univ Upsal 2021;2051. https://doi.org/10.33063/diva-442412.
[29] Pu B, Nan F, Zhu N, Yuan Y, Xie W. Ufngbm (1,1): a novel unbiased fractional grey Bernoulli model with Whale Optimization Algorithm and its application to electricity consumption forecasting in China. Energy Rep Nov. 2021;7:7405–23. https://doi.org/10.1016/j.egyr.2021.09.105.

[30] Saidi AYN, Ramli NA, Muhammad N, Awalin LJ. Power outage prediction by using logistic regression and decision tree. J. Phys.: Conf. Ser. Jul. 2021;1988(1):12039. https://doi.org/10.1088/1742-6596/1988/1/012039.
[31] Terrén-Serrano G, Martínez-Ramón M. Multi-layer wind velocity field visualization in infrared images of clouds for solar irradiance forecasting. Appl Energy Apr. 2021;288:116656. https://doi.org/10.1016/j.apenergy.2021.116656.
[32] Xu S, Wu W, Yang Y, Wang B, Wang X. Analytical solution of stochastic real-time dispatch incorporating wind power uncertainty characterized by Cauchy distribution. IET Renew Power Gener Jul. 2021;15(10):2286–301. https://doi.org/10.1049/rpg2.12163.
[33] Li C, Conejo AJ, Liu P, Omell BP, Siirola JD, Grossmann IE. Mixed-integer linear programming models and algorithms for generation and transmission expansion planning of power systems. Eur J Oper Res Mar. 2022;297(3):1071–82. https://doi.org/10.1016/j.ejor.2021.06.024.
[34] Pallonetto F, De Rosa M, D'Ettorre F, Finn DP. On the assessment and control optimisation of demand response programs in residential buildings. Renew Sustain Energy Rev Jul. 2020;127:109861. https://doi.org/10.1016/j.rser.2020.109861.
[35] Borunda M, Jaramillo OA, Reyes A, Ibargüengoytia PH. Bayesian networks in renewable energy systems: a bibliographical survey. Renew Sustain Energy Rev Sep. 2016;62:32–45. https://doi.org/10.1016/j.rser.2016.04.030.
[36] Amarasinghe K, Marino DL, Manic M. Deep neural networks for energy load forecasting. In: 2017 IEEE 26th international symposium on industrial electronics (ISIE), edinburgh, United Kingdom, Jun; 2017. p. 1483–8. https://doi.org/10.1109/ISIE.2017.8001465.
[37] Ahmad T, Zhang D, Huang C, Zhang H, Dai N, Song Y, et al. Artificial intelligence in sustainable energy industry: status quo, challenges and opportunities. J Clean Prod 2021;289:125834.
[38] Yılmaz B, Yıldırım R. Critical review of machine learning applications in perovskite solar research. Nano Energy 2021;80:105546. https://doi.org/10.1016/j.nanoen.2020.105546.
[39] Amasyali K, El-Gohary N. Machine learning for occupant-behavior-sensitive cooling energy consumption prediction in office buildings. Renew Sustain Energy Rev 2021;142:110714. https://doi.org/10.1016/j.rser.2021.110714.
[40] Xu X, Xu X, Jia Y, Xu Y, Xu Z, Xu Z, et al. A multi-agent reinforcement learning-based data-driven method for home energy management. IEEE Trans Smart Grid 2020;11:3201–11. https://doi.org/10.1109/TSG.2020.2971427.
[41] Du P. Ensemble machine learning-based wind forecasting to combine NWP output with data from weather station. IEEE Trans Sustain Energy 2019;10:2133–41. https://doi.org/10.1109/TSTE.2018.2880615.
[42] Hossain Lipu MS, Hannan MA, Hussain A, Saad MH, Ayob A, Uddin MN. Extreme learning machine model for state-of-charge estimation of lithium-ion battery using gravitational search algorithm. IEEE Trans Ind Appl 2019;55:4225–34. https://doi.org/10.1109/TIA.2019.2902532.
[43] Tayal A, Solanki A, Singh SP. Integrated frame work for identifying sustainable manufacturing layouts based on big data, machine learning, meta-heuristic and data envelopment analysis. Sustain Cities Soc 2020;62:102383. https://doi.org/10.1016/j.scs.2020.102383.
[44] King JE, Jupe SCE, Taylor PC. Network state-based algorithm selection for power flow management using machine learning. IEEE Trans Power Syst 2015;30:2657–64. https://doi.org/10.1109/TPWRS.2014.2361792.
[45] Zhang D, Li S, Sun M, O'Neill Z. An optimal and learning-based demand response and home energy management system. IEEE Trans Smart Grid 2016;7:1790–801. https://doi.org/10.1109/TSG.2016.2552169.
[46] Huang C, Zhang H, Song Y, Wang L, Ahmad T, Luo X. Demand response for industrial micro-grid considering photovoltaic power uncertainty and battery operational cost. IEEE Trans Smart Grid 2021:In press. https://doi.org/10.1109/TSG.2021.3052515.
[47] Susto GA, Schirru A, Pampuri S, McLoone S, Beghi A. Machine learning for predictive maintenance: a multiple classifier approach. IEEE Trans Ind Inf 2015;11:812–20. https://doi.org/10.1109/TII.2014.2349359.
[48] Ibrahim MS, Dong W, Yang Q. Machine learning driven smart electric power systems: current trends and new perspectives. Appl Energy 2020;272:115237. https://doi.org/10.1016/j.apenergy.2020.115237.
[49] Zhou Y, Zheng S, Liu Z, Wen T, Ding Z, Yan J, et al. Passive and active phase change materials integrated building energy systems with advanced machine-learning based climate-adaptive designs, intelligent operations, uncertainty-based analysis and optimisations: a state-of-the-art review. Renew Sustain Energy Rev 2020;130:109889. https://doi.org/10.1016/j.rser.2020.109889.
[50] Kellner T. Wind in the cloud? How the digital wind farm will make wind power 20 percent more efficient. GE Reports 2015:1–40. https://doi.org/http://www.gereports.com/post/119300678660/wind-in-the-cloud-how-the-digital-wind-farm-will/.
[51] Capgemini. Intelligent automation in energy and utilities. Capgemini Res Inst; 2019. p. 1–40. https://doi.org/https://www.capgemini.com/wp-content/uploads/2019/05/Digital-Report-%E2%80%93-Automation-in-Utilities-1.pdf.
[52] Pong PWT, Annaswamy AM, Kroposki B, Zhang Y, Rajagopal R, Zussman G, et al. Cyber-enabled grids: shaping future energy systems. Adv Appl Energy 2021;1:100003. https://doi.org/10.1016/j.adapen.2020.100003.
[53] Donnal JS, Leeb SB. Noncontact power meter. IEEE Sensor J 2015;15:1161–9. https://doi.org/10.1109/JSEN.2014.2359619.
[54] Tobin NP, McCormack M, O'Loughlin E, Tobin K. Remote current sensing on overhead power lines. In: IEE conf publ; 2001. p. 1. https://doi.org/10.1049/cp:20010727.
[55] Zhu K, Liu X, Pong PWT. On-site real-time current monitoring of three-phase three-core power distribution cables with magnetic sensing. Proc IEEE Sensors 2018;2018-Octob:1–4. https://doi.org/10.1109/ICSENS.2018.8589880.
[56] Power I, Society E. IEEE guide for application of optical instrument transformers for protective relaying IEEE guide for application of optical instrument transformers for protective relaying. EEE std C37241-2017. 2018. p. 1–50. https://doi.org/10.1109/IEEESTD.2018.8307291.
[57] Ghoddusi H, Creamer GG, Rafizadeh N. Machine learning in energy economics and finance: a review. Energy Econ 2019;81:709–27. https://doi.org/10.1016/j.eneco.2019.05.006.
[58] Wei J, De Luna P, Bengio Y, Aspuru-Guzik A, Sargent E. Use machine learning to find energy materials. Nature 2017;552:23–5. https://doi.org/10.1038/d41586-017-07820-6.
[60] Jain A, Ong SP, Hautier G, Chen W, Richards WD, Dacek S, et al. Commentary: the materials project: a materials genome approach to accelerating materials innovation. Apl Mater 2013;1:1–12. https://doi.org/10.1063/1.4812323.
[61] Chen C, Zuo Y, Ye W, Li X, Deng Z, Ong SP. A critical review of machine learning of energy materials. Adv Energy Mater 2020;10:1–36. https://doi.org/10.1002/aenm.201903242.
[62] Zhang L, Wen J, Li Y, Chen J, Ye Y, Fu Y, et al. A review of machine learning in building load prediction. Appl Energy 2021;285:116452. https://doi.org/10.1016/j.apenergy.2021.116452.
[63] Ahmad R, Alsmadi I. Machine learning approaches to IoT security: a systematic literature review. Internet of Things 2021;14:100365. https://doi.org/10.1016/j.iot.2021.100365.
[64] Kang Y, Li L, Li B. Recent progress on discovery and properties prediction of energy materials: simple machine learning meets complex quantum chemistry. J Energy Chem 2021;54:72–88. https://doi.org/10.1016/j.jechem.2020.05.044.
[65] Turetskyy A, Wessel J, Herrmann C, Thiede S. Battery production design using multi-output machine learning models. Energy Storage Mater 2021:1–24.
[66] Rahnama A, Zepon G, Sridhar S. Machine learning based prediction of metal hydrides for hydrogen storage, part I: prediction of hydrogen weight percent. Int J Hydrogen Energy 2019;44:7337–44. https://doi.org/10.1016/j.ijhydene.2019.01.261.
[67] Zhou Y, Zheng S, Zhang G. A state-of-the-art-review on phase change materials integrated cooling systems for deterministic parametrical analysis, stochastic uncertainty-based design, single and multi-objective optimisations with machine learning applications. Energy Build 2020;220:110013. https://doi.org/10.1016/j.enbuild.2020.110013.
[68] Gao T, Lu W. Machine learning toward advanced energy storage devices and systems. iScience 2021;24:101936. https://doi.org/10.1016/j.isci.2020.101936.
[69] Wang Y, Liu J, Han Y. Production capacity prediction of hydropower industries for energy optimization: evidence based on novel extreme learning machine integrating Monte Carlo. J Clean Prod 2020;272:122824. https://doi.org/10.1016/j.jclepro.2020.122824.
[70] Seferlis P, Varbanov PS, Papadopoulos AI, Chin HH, Klemeš JJ. Sustainable design, integration, and operation for energy high-performance process systems. Energy 2021;224:120158. https://doi.org/10.1016/j.energy.2021.120158.
[71] Narciso DAC, Martins FG. Application of machine learning tools for energy efficiency in industry: a review. Energy Rep 2020;6:1181–99. https://doi.org/10.1016/j.egyr.2020.04.035.
[72] Ikeda S, Nagai T. A novel optimization method combining metaheuristics and machine learning for daily optimal operations in building energy and storage systems. Appl Energy 2021;289:116716. https://doi.org/10.1016/j.apenergy.2021.116716.
[73] Ahmad T, Chen H. A review on machine learning forecasting growth trends and their real-time applications in different energy systems. Sustain Cities Soc 2020;54:102010. https://doi.org/10.1016/j.scs.2019.102010.
[74] Ahmad T, Zhang D. Using the internet of things in smart energy systems and networks. Sustain Cities Soc 2021;68:102783. https://doi.org/10.1016/j.scs.2021.102783.
[75] Ali SS, Choi BJ. State-of-the-art artificial intelligence techniques for distributed smart grids: a review. Electron 2020;9:1–28. https://doi.org/10.3390/electronics9061030.
[76] Tayal D. Achieving high renewable energy penetration in Western Australia using data digitisation and machine learning. Renew Sustain Energy Rev 2017;80:1537–43. https://doi.org/10.1016/j.rser.2017.07.040.
[77] Advances S. Using AI to improve energy and resource efficiency in various industries. Eur Community 2021. https://doi.org/https://cordis.europa.eu/article/id/415798-using-ai-to-improve-energy-and-resource-efficiency-in-various-industries, . [Accessed 3 March 2021].
[78] Tomazzoli C, Scannapieco S. Machine learning for energy efficiency: automatic detection of electric loads from power consumption. In: 2017 AEIT int annu conf infrastructures energy ICT oppor foster innov AEIT 2017 2017; 2017-January. p. 1–6. https://doi.org/10.23919/AEIT.2017.8240544.
[79] Paris EH, Washington SR, Chui M, Francisco S, London TA. Artificial intelligence the next digital frontier ? McKinsey Co; 2017. p. 1–80. Sree Ramas.
[80] Burlig F, Knittel C, Rapson D, Reguant M, Wolfram C. Machine learning from schools about energy efficiency. J Assoc Environ Resour Econ 2020;7:1181–217. https://doi.org/10.1086/710606.
[81] Owoyele O, Pal P. A novel machine learning-based optimization algorithm (ActivO) for accelerating simulation-driven engine design. Appl Energy 2021;285:116455. https://doi.org/10.1016/j.apenergy.2021.116455.
[82] Zhou Y, Zheng S. Climate adaptive optimal design of an aerogel glazing system with the integration of a heuristic teaching-learning-based algorithm in machine

learning-based optimization. Renew Energy 2020;153:375–91. https://doi.org/10.1016/j.renene.2020.01.133.
[83] Al-Shahri OA, Ismail FB, Hannan MA, Lipu MSH, Al-Shetwi AQ, Begum RA, et al. Solar photovoltaic energy optimization methods, challenges and issues: a comprehensive review. J Clean Prod 2021;284:125465. https://doi.org/10.1016/j.jclepro.2020.125465.
[84] Kong X, Li C, Zheng F, Wang C. Improved deep belief network for short-term load forecasting considering demand-side management. IEEE Trans Power Syst 2020;35:1531–8. https://doi.org/10.1109/TPWRS.2019.2943972.
[85] Lopez KL, Gagne C, Gardner MA. Demand-side management using deep learning for smart charging of electric vehicles. IEEE Trans Smart Grid 2019;10:2683–91. https://doi.org/10.1109/TSG.2018.2808247.
[86] Mahmud K, Khan B, Ravishankar J, Ahmadi A, Siano P. An internet of energy framework with distributed energy resources, prosumers and small-scale virtual power plants: an overview. Renew Sustain Energy Rev 2020;127:109840. https://doi.org/10.1016/j.rser.2020.109840.
[87] IRENA. Innovation landscape for a renewable-powered futureSOLUTION V; 2019. p. 1–6. /Publications/2019/Feb/Innovation-Landscape-for-a-Renewable-Powered-Future https://doi.org/https://www.irena.org/publications/2019/Feb/Innovation-landscape-for-a-renewable-powered-future.
[88] Vázquez-Canteli JR, Nagy Z. Reinforcement learning for demand response: a review of algorithms and modeling techniques. Appl Energy 2019;235:1072–89. https://doi.org/10.1016/j.apenergy.2018.11.002.
[89] Haq EU, Lyu X, Jia Y, Hua M, Ahmad F. Forecasting household electric appliances consumption and peak demand based on hybrid machine learning approach. Energy Rep 2020;6:1099–105. https://doi.org/10.1016/j.egyr.2020.11.071.
[90] Wen L, Zhou K, Li J, Wang S. Modified deep learning and reinforcement learning for an incentive-based demand response model. Energy 2020;205:118019. https://doi.org/10.1016/j.energy.2020.118019.
[91] Ponnaganti P, Pillai JR, Bak-Jensen B. Opportunities and challenges of demand response in active distribution networks, vol. 7. Wiley Interdiscip Rev Energy Environ; 2018. p. 1–16. https://doi.org/10.1002/wene.271.
[92] Ghobakhloo M. Industry 4.0, digitization, and opportunities for sustainability. J Clean Prod 2020;252:119869. https://doi.org/10.1016/j.jclepro.2019.119869.
[93] Hu M, Chen H, Shen L, Li G, Guo Y, Li H, et al. A machine learning bayesian network for refrigerant charge faults of variable refrigerant flow air conditioning system. Energy Build 2018;158:668–76. https://doi.org/10.1016/j.enbuild.2017.10.012.
[94] Kück M, Freitag M. Forecasting of customer demands for production planning by local k-nearest neighbor models. Int J Prod Econ 2021;231:107837. https://doi.org/10.1016/j.ijpe.2020.107837.
[95] Yang L, He M, Zhang J, Vittal V. Support-vector-machine-enhanced markov model for short-term wind power forecast. IEEE Trans Sustain Energy 2015;6:791–9. https://doi.org/10.1109/TSTE.2015.2406814.
[96] Kawase H, Mori Y, Hasegawa H, Sato KC. Dynamic router performance control utilizing support vector machines for energy consumption reduction. IEEE Trans Netw Serv Manag 2016;13:860–70. https://doi.org/10.1109/TNSM.2016.2605640.
[97] Perera ATD, Kamalaruban P. Applications of reinforcement learning in energy systems. Renew Sustain Energy Rev 2021;137:110618. https://doi.org/10.1016/j.rser.2020.110618.
[98] Zhang X, Li Y, Lu S, Hamann HF, Hodge BM, Lehman B. A solar time based analog ensemble method for regional solar power forecasting. IEEE Trans Sustain Energy 2019;10:268–79. https://doi.org/10.1109/TSTE.2018.2832634.
[99] Han Y, Xiu L, Wang Z, Chen Q, Tan S. Artificial neural networks controlled fast valving in a power generation plant. IEEE Trans Neural Network 1997;8:373–89. https://doi.org/10.1109/72.557689.
[100] Dai X, Liu J, Zhang X. A review of studies applying machine learning models to predict occupancy and window-opening behaviours in smart buildings. Energy Build 2020;223:110159. https://doi.org/10.1016/j.enbuild.2020.110159.
[101] Ray PK, Mohanty SR, Kishor N, Catalao JPS. Optimal feature and decision tree-based classification of power quality disturbances in distributed generation systems. IEEE Trans Sustain Energy 2014;5:200–8. https://doi.org/10.1109/TSTE.2013.2278865.
[102] Ramos V, Giannini G, Calheiros-Cabral T, Rosa-Santos P, Taveira-Pinto F. Legal framework of marine renewable energy: a review for the Atlantic region of Europe. Renew Sustain Energy Rev 2021;137:110608. https://doi.org/10.1016/j.rser.2020.110608.
[103] Liukkonen M, Hiltunen T. Adaptive monitoring of emissions in energy boilers using self-organizing maps: an application to a biomass-fired CFB (circulating fluidized bed). Energy 2014;73:443–52. https://doi.org/10.1016/j.energy.2014.06.034.
[104] Saraswat SK, Digalwar AK. Evaluation of energy alternatives for sustainable development of energy sector in India: an integrated Shannon's entropy fuzzy multi-criteria decision approach. Renew Energy 2021;171:58–74. https://doi.org/10.1016/j.renene.2021.02.068.
[105] Baharoon DA, Rahman HA, Fadhl SO. Publics' knowledge, attitudes and behavioral toward the use of solar energy in Yemen power sector. Renew Sustain Energy Rev 2016;60:498–515. https://doi.org/10.1016/j.rser.2015.12.110.
[106] Kotsiopoulos T, Sarigiannidis P, Ioannidis D, Tzovaras D. Machine learning and deep learning in smart manufacturing: the smart grid paradigm. Comput Sci Rev 2021;40:100341. https://doi.org/10.1016/j.cosrev.2020.100341.
[107] Qu YJ, Ming XG, Liu ZW, Zhang XY, Hou ZT. Smart manufacturing systems: state of the art and future trends. Int J Adv Manuf Technol 2019;103:3751–68. https://doi.org/10.1007/s00170-019-03754-7.
[108] Lee J, Davari H, Singh J, Pandhare V. Industrial Artificial Intelligence for industry 4.0-based manufacturing systems. Manuf Lett 2018;18:20–3. https://doi.org/10.1016/j.mfglet.2018.09.002.
[109] Wang H, Lei Z, Zhang X, Zhou B, Peng J. A review of deep learning for renewable energy forecasting. Energy Convers Manag 2019;198. https://doi.org/10.1016/j.enconman.2019.111799.
[110] Ahmad T, Zhang H, Yan B. A review on renewable energy and electricity requirement forecasting models for smart grid and buildings. Sustain Cities Soc 2020;55. https://doi.org/10.1016/j.scs.2020.102052.
[111] Tavana M, Shaabani A, Santos-Arteaga FJ, Vanani IR. A review of uncertain decision-making methods in energy management using text mining and data analytics. Energies 2020;13:1–13. https://doi.org/10.3390/en13153947.
[112] Nisar A, Ruiz F, Palacios M. Organisational learning, strategic rigidity and technology adoption: implications for electric utilities and renewable energy firms. Renew Sustain Energy Rev 2013;22:438–45. https://doi.org/10.1016/j.rser.2013.01.039.
[113] Gupta V, Bali R. Artificial intelligence in the energy world—getting the act together. 2021. https://doi.org/10.1007/978-981-15-8297-4_63. 793–8.
[114] Nikkei. AI to propel wind farm efficiency in Japan. AsiaNikkeiCom/Business/Companies/AI-to-Propel-Wind-Farm-Efficiency-in-Japan 2017. https://doi.org/https://asia. nikkei.com/Business/AI-to-propel-wind-farmefficiency- in-Japan.
[115] Company M. Performance management and optimization for the electricpower industry. McKinsey Co; 2019. p. 1–24 [https://doi.org/www.mckinsey.com/solutions/utilityx].
[116] Mohanta BK, Jena D, Satapathy U, Patnaik S. Survey on IoT security: challenges and solution using machine learning, artificial intelligence and blockchain technology. Internet of Things 2020;11:100227. https://doi.org/10.1016/j.iot.2020.100227.
[117] Henzelmann Torsten, Hammermeister F. Artificial intelligence: a smart move for utilities. Rol Berger ThinkAct 2018.
[118] Heffron R, Körner MF, Wagner J, Weibelzahl M, Fridgen G. Industrial demand-side flexibility: a key element of a just energy transition and industrial development. Appl Energy 2020;269:115026. https://doi.org/10.1016/j.apenergy.2020.115026.
[119] Verhees B, Raven R, Kern F, Smith A. The role of policy in shielding, nurturing and enabling offshore wind in The Netherlands (1973-2013). Renew Sustain Energy Rev 2015;47:816–29. https://doi.org/10.1016/j.rser.2015.02.036.
[120] Magazzino C, Mele M, Schneider N. A machine learning approach on the relationship among solar and wind energy production, coal consumption, GDP, and CO2 emissions. Renew Energy 2021;167:99–115. https://doi.org/10.1016/j.renene.2020.11.050.
[121] Gorjian S, Sharon H, Ebadi H, Kant K, Scavo FB, Tina GM. Recent technical advancements, economics and environmental impacts of floating photovoltaic solar energy conversion systems. J Clean Prod 2021;278:124285. https://doi.org/10.1016/j.jclepro.2020.124285.
[122] Elshawi R, Sakr S. Big data systems meet machine learning challenges: towards big data science as a service. Big Data Res 2018;14:1–11.
[123] Onile AE, Machlev R, Petlenkov E, Levron Y, Belikov J. Uses of the digital twins concept for energy services, intelligent recommendation systems, and demand side management: a review. Energy Rep 2021;7:997–1015. https://doi.org/10.1016/j.egyr.2021.01.090.
[124] Zhou Y, Liu Y, Wang D, Liu X, Wang Y. A review on global solar radiation prediction with machine learning models in a comprehensive perspective. Energy Convers Manag 2021;235:113960. https://doi.org/10.1016/j.enconman.2021.113960.
[125] Dogan A, Birant D. Machine learning and data mining in manufacturing. Expert Syst Appl 2021;166:114060. https://doi.org/10.1016/j.eswa.2020.114060.
[126] Liu Y, Guo B, Zou X, Li Y, Shi S. Machine learning assisted materials design and discovery for rechargeable batteries. Energy Storage Mater 2020;31:434–50. https://doi.org/10.1016/j.ensm.2020.06.033.
[127] Oracle Oracle. Big data for the enterprise. Oracle White Pap 2012:16.
[128] Zhang Y, Huang T, Bompard EF. Big data analytics in smart grids: a review. Energy Informatics 2018;1:1–8. https://doi.org/10.1186/s42162-018-0007-5.
[129] Bhattarai BP, Paudyal S, Luo Y, Mohanpurkar M, Cheung K, Tonkoski R, et al. Big data analytics in smart grids: state-of-theart, challenges, opportunities, and future directions. IET Smart Grid 2019;2:141–54. https://doi.org/10.1049/iet-stg.2018.0261.
[130] Gu B, Sheng VS. A robust regularization path algorithm for ν-support vector classification. IEEE Transact Neural Networks Learn Syst 2017;28:1241–8. https://doi.org/10.1109/TNNLS.2016.2527796.
[131] GTMResearch. Sas. The soft grid 2013-2020 : big data & utility analytics for smart grid. 2013. p. 1–15.
[132] Zhou L, Pan S, Wang J, Vasilakos AV. Machine learning on big data: opportunities and challenges. Neurocomputing 2017;237:350–61. https://doi.org/10.1016/j.neucom.2017.01.026.
[133] Bayram IS, Saffouri F, Koc M. Generation, analysis, and applications of high resolution electricity load profiles in Qatar. J Clean Prod 2018;183:527–43. https://doi.org/10.1016/j.jclepro.2018.02.084.
[134] Patnam BSK, Pindoriya NM. Demand response in consumer-Centric electricity market: mathematical models and optimization problems. Elec Power Syst Res 2020;193:106923. https://doi.org/10.1016/j.epsr.2020.106923.
[135] Fathi S, Srinivasan R, Fenner A, Fathi S. Machine learning applications in urban building energy performance forecasting: a systematic review. Renew Sustain Energy Rev 2020;133:110287. https://doi.org/10.1016/j.rser.2020.110287.
[136] Seyedzadeh S, Pour Rahimian F, Oliver S, Rodriguez S, Glesk I. Machine learning modelling for predicting non-domestic buildings energy performance: a model to

support deep energy retrofit decision-making. Appl Energy 2020;279:115908. https://doi.org/10.1016/j.apenergy.2020.115908.
[137] Said Z, Hachicha AA, Aberoumand S, Yousef BAA, Sayed ET, Bellos E. Recent advances on nanofluids for low to medium temperature solar collectors: energy, exergy, economic analysis and environmental impact. Prog Energy Combust Sci 2021;84:100898. https://doi.org/10.1016/j.pecs.2020.100898.
[138] Sun W, Zhang C. Analysis and forecasting of the carbon price using multi—resolution singular value decomposition and extreme learning machine optimized by adaptive whale optimization algorithm. Appl Energy 2018:1354–71. https://doi.org/10.1016/j.apenergy.2018.09.118.
[139] Sensfuß F, Ragwitz M, Genoese M. The merit-order effect: a detailed analysis of the price effect of renewable electricity generation on spot market prices in Germany. Energy Pol 2008;36:3086–94. https://doi.org/10.1016/j.enpol.2008.03.035.
[140] Kell AJM, McGough AS, Forshaw M. The impact of online machine-learning methods on long-term investment decisions and generator utilization in electricity markets. Sustain Comput Informatics Syst 2020;30:100532.
[141] Freitas Gomes IS, Perez Y, Suomalainen E. Coupling small batteries and PV generation: a review. Renew Sustain Energy Rev 2020;126:109835. https://doi.org/10.1016/j.rser.2020.109835.
[142] Strielkowski W. Consumers, prosumers, and the smart grids. Soc Impacts Smart Grids 2020:191–238. https://doi.org/10.1016/b978-0-12-817770-9.00006-7.
[143] GuidoPerboli; Ehsan Arabnezhad. A Machine Learning-based DSS for mid and long-term company crisis prediction. Expert Syst Appl n.d.;174:114758.
[144] Strielkowski W. Traditional power markets and an evolution to smart grids. Soc Impacts Smart Grids 2020:9–54. https://doi.org/10.1016/b978-0-12-817770-9.00002-x.
[145] Babuta A, Gupta B, Kumar A, Ganguli S. Power and energy measurement devices: a review, comparison, discussion, and the future of research. Meas J Int Meas Confed 2021;172:108961. https://doi.org/10.1016/j.measurement.2020.108961.
[146] du Plessis AA, Strauss JM, Rix AJ. Short-term solar power forecasting: investigating the ability of deep learning models to capture low-level utility-scale Photovoltaic system behaviour. Appl Energy 2021;285:116395. https://doi.org/10.1016/j.apenergy.2020.116395.
[147] Park C, Heo WG. Review of the changing electricity industry value chain in the ICT convergence era. J Clean Prod 2020;258:120743. https://doi.org/10.1016/j.jclepro.2020.120743.
[148] Ferrando M, Causone F, Hong T, Chen Y. Urban building energy modeling (UBEM) tools: a state-of-the-art review of bottom-up physics-based approaches. Sustain Cities Soc 2020;62:102408. https://doi.org/10.1016/j.scs.2020.102408.
[149] Robinson D, Haldi F, Kämpf J, Leroux P, Perez D, Rasheed A, et al. Citysim: comprehensive micro-simulation of resource flows for sustainable urban planning. IBPSA 2009 - Int Build Perform Simul Assoc 2009 2009. 1083–90.
[150] Nouvel R, Brassel K-H, Bruse M, Duminil E, Coors V, Eicker U, et al. SIMSTADT, a new workflow-driven urban energy simulation platform for CityGML city models. In: CISBAT int conf; 2015. p. 889–94.
[151] Agugiaro G, Benner J, Cipriano P, Nouvel R. The Energy Application Domain Extension for CityGML: enhancing interoperability for urban energy simulations. Open Geospatial Data, Softw Stand 2018;3:1–2. https://doi.org/10.1186/s40965-018-0042-y.
[152] CityGML. Open geospatial consort. 2021 [https://doi.org/https://www.citygml.org/].
[153] MIT. Sustainable design lab. Umi - urban modeling interface. 2013 [https://doi.org/http://web.mit.edu/sustainabledesignlab/projects/umi/index.html].
[154] Hong T, Chen Y, Piette Xl MA. City 2017 [https://doi.org/https://citybes.lbl.gov/].
[155] Fonseca JA, Nguyen TA, Schlueter A, Marechal F. City Energy Analyst (CEA): integrated framework for analysis and optimization of building energy systems in neighborhoods and city districts. Energy Build 2016;113:202–26. https://doi.org/10.1016/j.enbuild.2015.11.055.
[156] Remmen P, Lauster M, Mans M, Fuchs M, Osterhage T, Müller D. TEASER: an open tool for urban energy modelling of building stocks. J Build Perform Simul 2018;11:84–98. https://doi.org/10.1080/19401493.2017.1283539.
[157] Menghi R, Papetti A, Germani M, Marconi M. Energy efficiency of manufacturing systems: a review of energy assessment methods and tools. J Clean Prod 2019;240:118276. https://doi.org/10.1016/j.jclepro.2019.118276.
[158] Lund H. Tool: the energyplan energy system analysis model. Renew Energy Syst A Smart Energy Syst Approach to Choice Model 100% Renew Solut Second 2014:53–78. https://doi.org/10.1016/B978-0-12-410423-5.00004-3.
[159] Ferrari S, Zagarella F, Caputo P, Bonomolo M. Assessment of tools for urban energy planning. Energy 2019;176:544–51. https://doi.org/10.1016/j.energy.2019.04.054.
[160] Yao L, Yang B, Cui H, Zhuang J, Ye J, Xue J. Challenges and progresses of energy storage technology and its application in power systems. J Mod Power Syst Clean Energy 2016;4:519–28. https://doi.org/10.1007/s40565-016-0248-x.
[161] Lu Y, Zheng X6G. A survey on technologies, scenarios, challenges, and the related issues. J Ind Inf Integr 2020;19:100158. https://doi.org/10.1016/j.jii.2020.100158.
[162] Oconnell N, Pinson P, Madsen H, Omalley M. Benefits and challenges of electrical demand response: a critical review. Renew Sustain Energy Rev 2014;39:686–99. https://doi.org/10.1016/j.rser.2014.07.098.
[163] Artificial Intelligence and Big Data Innovation Landscape Brief. Irea 2019. IRENA; 2019. p. 1–24. https://doi.org/https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2019/Sep/IRENA_AI_Big_Data_2019.pdf?la=en&hash=9A003F48B639B810237FEEAF61D47C74F8D8F07F.
[164] Nižetić S, Šolić P, López-de-Ipiña González-de-Artaza D, Patrono L. Internet of Things (IoT): opportunities, issues and challenges towards a smart and sustainable future. J Clean Prod 2020;274:122877. https://doi.org/10.1016/j.jclepro.2020.122877.
[165] Feng C, Wang Y, Chen Q, Ding Y, Strbac G, Kang C. Smart grid encounters edge computing: opportunities and applications. Adv Appl Energy 2021;1:100006. https://doi.org/10.1016/j.adapen.2020.100006.
[166] Amin AZ. Global energy transformation. Int Renew Energy Agency 2018:1–76. https://doi.org/10.1057/9780230244092.
[167] You NZF. New York State's 100% renewable electricity transition planning under uncertainty using a data-driven multistage adaptive robust optimization approach with machine-learning. Adv Appl Energy n.d.;Available:100019. https://doi.org/https://doi.org/10.1016/j.adapen.2021.100019.
[168] Vasudevan RK, Choudhary K, Mehta A, Smith R, Kusne G, Tavazza F, et al. Materials science in the artificial intelligence age: high-throughput library generation, machine learning, and a pathway from correlations to the underpinning physics. MRS Commun 2019;9:821–38. https://doi.org/10.1557/mrc.2019.95.
[169] Huang JS, Liew JX, Ademiloye AS, Liew KM. Artificial intelligence in materials modeling and design. Arch Comput Methods Eng 2020. https://doi.org/10.1007/s11831-020-09506-1.
[170] Zhou M, Gallegos A, Liu K, Dai S, Wu J. Insights from machine learning of carbon electrodes for electric double layer capacitors. Carbon N Y 2020;157:147–52. https://doi.org/10.1016/j.carbon.2019.08.090.
[171] Orupattur NV, Mushrif SH, Prasad V. Catalytic materials and chemistry development using a synergistic combination of machine learning and ab initio methods. Comput Mater Sci 2020;174:109474. https://doi.org/10.1016/j.commatsci.2019.109474.
[172] Yang W, Fidelis TT, Sun WH. Machine learning in catalysis, from proposal to practicing. ACS Omega 2020;5:83–8. https://doi.org/10.1021/acsomega.9b03673.
[173] Zhang L, He M, Shao S. Machine learning for halide perovskite materials. Nano Energy 2020;78:105380. https://doi.org/10.1016/j.nanoen.2020.105380.
[174] Odabaşı Ç, Yıldırım R. Machine learning analysis on stability of perovskite solar cells. Sol Energy Mater Sol Cells 2020;205:110284. https://doi.org/10.1016/j.solmat.2019.110284.
[175] Shi Z, Yang W, Deng X, Cai C, Yan Y, Liang H, et al. Machine-learning-assisted high-throughput computational screening of high performance metal-organic frameworks. Mol Syst Des Eng 2020;5:725–42. https://doi.org/10.1039/d0me00005a.
[176] He Y, Cubuk ED, Allendorf MD, Reed EJ. Metallic metal-organic frameworks predicted by the combination of machine learning methods and ab initio calculations. J Phys Chem Lett 2018;9:4562–9. https://doi.org/10.1021/acs.jpclett.8b01707.
[177] Toyao T, Suzuki K, Kikuchi S, Takakusagi S, Shimizu KI, Takigawa I. Toward effective utilization of methane: machine learning prediction of adsorption energies on metal alloys. J Phys Chem C 2018;122:8315–26. https://doi.org/10.1021/acs.jpcc.7b12670.
[178] Vazquez S, Lukic S, Galvan E, Franquelo LG, Carrasco JM, Leon JI. Recent advances on energy storage systems. In: IECON proc (industrial electron conf; 2011. p. 4636–40. https://doi.org/10.1109/IECON.2011.6120075.
[179] Intelligence A. Spotlight: artificial intelligence. Using artificial intelligence to advance the state of multiple industries. September 2019. 2019. https://doi.org/https://www.energy.gov/sites/prod/files/2019/10/f67/2019-09-30_Spotlight-Artificial_Intelligence_0.pdf.
[180] Shaukat N, Ali SM, Mehmood CA, Khan B, Jawad M, Farid U, et al. A survey on consumers empowerment, communication technologies, and renewable generation penetration within Smart Grid. Renew Sustain Energy Rev 2018;81:1453–75. https://doi.org/10.1016/j.rser.2017.05.208.
[181] Dehghanpour K, Hashem Nehrir M, Sheppard JW, Kelly NC. Agent-based modeling of retail electrical energy markets with demand response. IEEE Trans Smart Grid 2018;9:3465–75. https://doi.org/10.1109/TSG.2016.2631453.
[182] Zhao W, Zhang H, Zheng J, Dai Y, Huang L, Shang W, et al. A point prediction method based automatic machine learning for day-ahead power output of multi-region photovoltaic plants. Energy 2021;223:120026. https://doi.org/10.1016/j.energy.2021.120026.
[183] Khodayar M, Mohammadi S, Khodayar ME, Wang J, Liu G. Convolutional graph autoencoder: a generative deep neural network for probabilistic spatio-temporal solar irradiance forecasting. IEEE Trans Sustain Energy 2020;11:571–83. https://doi.org/10.1109/TSTE.2019.2897688.
[184] Guevara E, Babonneau F, Homem-de-Mello T, Moret S. A machine learning and distributionally robust optimization framework for strategic energy planning under uncertainty. Appl Energy 2020;271:115005. https://doi.org/10.1016/j.apenergy.2020.115005.
[185] Hu L, Che XL, Zheng SQ. Online system for grid resource monitoring and machine learning-based prediction. IEEE Trans Parallel Distr Syst 2012;23:134–45. https://doi.org/10.1109/TPDS.2011.108.
[186] Wen S, Zhang C, Lan H, Xu Y, Tang Y, Huang Y. A hybrid ensemble model for interval prediction of solar power output in ship onboard power systems. IEEE Trans Sustain Energy 2021;12:14–24. https://doi.org/10.1109/TSTE.2019.2963270.
[187] Paris EH, Washington SR, Chui M, Francisco S, London TA. Artificial intelligence the next digital frontier. McKinsey Co. Sree Ramas 2017:1–80.
[188] IRENA - International Renewable Energy Agency. REmap 2030 - a renewable energy roadmap. IRENA; 2014. p. 1–56. https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2014/IRENA_REmap_summary_findings_2014.pdf.