A Review of Federated Learning in Energy Systems

Xu Cheng
Section of Energy Markets
Smart Innovation Norway
Halden, Norway
xu.cheng@ieee.org

Chendan Li
Department of Marine Technology
Norwegian University of Science and Technology
Trondheim, Norway
chendan.li@ntnu.no

Xiufeng Liu
Department of Technology,
Management and Economics
Technical University of Denmark
Produktionstorvet, Denmark
xiuli@dtu.dk

Abstract—**With increasing concerns for data privacy and ownership, recent years have witnessed a paradigm shift in machine learning (ML). An emerging paradigm, federated learning (FL), has gained great attention and has become a novel design for machine learning implementations. FL enables the ML model training at data silos under the coordination of a central server, eliminating communication overhead and without sharing raw data. In this paper, we conduct a review of the FL paradigm and, in particular, compare the types, the network structures, and the global model aggregation methods. Then, we conducted a comprehensive review of FL applications in the energy domain (refer to the smart grid in this paper). We provide a thematic classification of FL to address a variety of energy-related problems, including demand response, identification, prediction, and federated optimizations. We describe the taxonomy in detail and conclude with a discussion of various aspects, including challenges, opportunities, and limitations in its energy informatics applications, such as energy system modeling and design, privacy, and evolution.**

Index Terms—**Review, Federated Learning, Energy sector, Distributed learning, Privacy and security**

I. INTRODUCTION

Today the demand for energy is increasing rapidly in the global, due to population and economic growth. The building sector accounts for 40% of total energy consumption and 60% of electricity consumption [1]. Distributed energy resources are a viable solution to address this problem by integrating distributed energy systems at users. Distributed energy systems have the characteristics of high efficiency, low loss, low pollution, and flexible operations. On the other hand, for electricity supply, smart grids have intensively adopted the Internet of Things (IoT), smart meters, and advanced communication networks (e.g. 5G) and data management systems, which form an advanced energy management infrastructure. Smart grids and distributed energy systems together form a giant energy of the Internet (IoE) to provide stable and sustainable energy to end users. However, several prominent issues arise with the provision of various energy services, the management of massive data, and the offloading of computing workloads from the cloud. First, data privacy becomes a major concern as customer data is often used to train data-driven models, e.g., for detecting malicious nodes in IoE. Therefore, it has to access private customer information, such as consumption profiles and habits, to train the model. Second, data owners may not be willing to share their data to a centralized server for model training, mainly due to privacy,

commercial competition, and technical barrier reasons [2]. At the legal level, governments or organizations around the world are also increasingly committed to data privacy protection. For example, the European Union has enforced the General Data Protection Regulation (GDPR) since 2017 [3]. Third, the computing resource requirements for traditional centralized model training are relatively high due to the management of large amounts of data sets. Last, data transmission over the network is vulnerable to cyberattacks and introduces overhead. Therefore, there is a pressing need for a feasible solution to address these issues.

Federated learning (FL) is an ideal paradigm, which was first introduced by Google in 2016 [4]. FL is essentially a distributed learning approach, where multiple distributed clients train machine learning models separately using their own local data under the coordination of a central server, then the server aggregates the trained models into a final global model. Training is an iterative process that involves updating local models and aggregating the global model until the model converges or the training reaches the predefined number of interactions. FL is a major shift from centralized and expensive machine learning to a distributed manner that can use many distributed computing resources. This learning paradigm improves data privacy, as the raw data remain on the local device, eliminating network overhead as only the model updates are exchanged. In recent years, FL has been used successfully in industries related to IoT and edge computing, mainly for privacy protection, but it is attracting more and more interest. Some literature reviews were performed not only on FL technology itself, such as the different FL architectures and their learning settings, e.g., [5]–[10], but also its applications, e.g., [11]–[15]. However, there is still no systematic review of FL applications in energy. This is likely due to more research efforts in areas directly related to the protection of personal information privacy, such as medicine, healthcare, and insurance. On the contrary, the literature on FL in energy is much fewer and all of them emerged within the last two years. However, we have seen that growth is fast and its applications go beyond addressing data privacy, which is raising great research attention in the energy community. Therefore, we believe that it is necessary to conduct a systematic review. In this paper, we will classify and summarize the different solutions based on FL and their applications in the energy sector. Moreover, unlike other reviews

of FL technologies which focus on different architectures and complementary technologies to enhance FL, we pay special attention to different aggregate algorithms to generate the global model. This review can help advance understanding of FL technology and promote its use cases in the energy sector.

In summary, this paper makes the following contributions.

- We list the taxonomy of FL technology, classify FL frameworks according to data partitioning methods and network structures, and, in particular, summarize the mainstream aggregation algorithms for generating the global model.
- We review FL use cases in the smart grid domain, identify their technological features, and provide an overview of current FL in the energy domain.
- We discuss the challenges and opportunities for the FL technology itself and for applications in energy and suggest possible solutions to shed light on the wider adoption of FL in energy.

The remainder of this paper is structured as follows. Section II gives an overview of the FL paradigm, and surveys the methods for global model generation; Section III reviews the use cases of FL in energy systems; Section IV discusses the challenges, opportunities, and limitations of FL in energy; Section V concludes the paper and presents the future research directions.

II. FEDERATED LEARNING CONCEPT

A. Overview

FL provides a novel machine learning solution with distributed and enhanced privacy capabilities for many energy applications, which has the potential to shape current energy systems [16]. With growing concerns about privacy issues, FL is particularly attractive for constructing a distributed model by pushing machine learning functions to the edge of the energy system when the data reside. Eventually, the data of each energy system is never directly shared with others under the FL framework, while supporting collaborative learning of shared global models, which benefits both grid operators and energy users in terms of network resource savings and enhanced privacy. As such, FL is a powerful alternative to traditional centralized machine learning approaches.

Figure 1 illustrates the concept of FL, where there are two entities: clients and a central server. The local models can be the end users of an energy system, and the central server is used to aggregate the local models trained by the clients. The general FL process comprises the following three steps:

1) **System initialization and device selection**: Most recent work assumes that all clients participate in each training round in the FL process. However, in each training round, only a certain number of clients are selected in practice.
2) **Local training**: When training starts, the central server initializes a global model and distributes it to all selected clients. The clients i then start local training using their
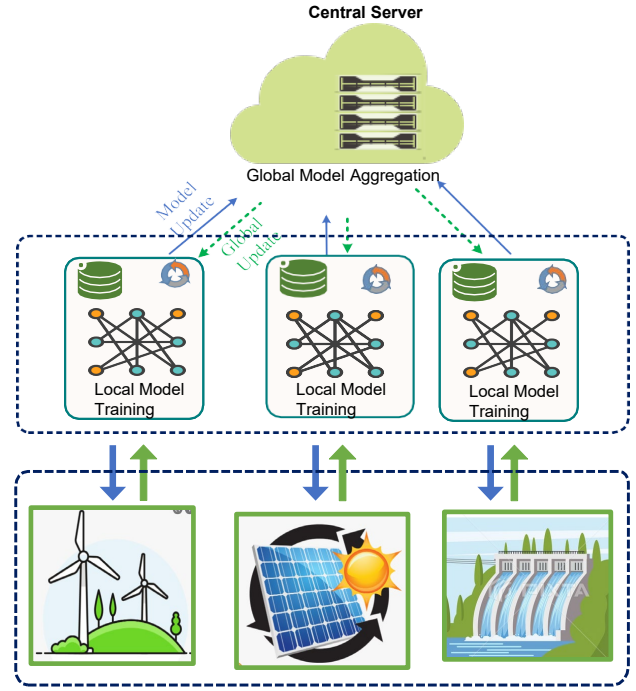

Fig. 1. Overview of federated learning applied in an energy system

own local data d_i and calculate the gradient ω_i by minimizing a loss function $F(\omega_i)$:

$$\omega_i^* = argminF(\omega_i) \qquad (1)$$

Note that the loss function varies between different FL algorithms depending on the relationship between the input and output pairs.

3) **Model aggregation**: After collecting all model updates from the clients, the server will calculate a new global model as :

$$\omega_S = \frac{1}{\sum_{i \in N} |d_i|} \sum_{i=1}^{N} |d_i| w_i \qquad (2)$$

where ω_S represents the global model. After deriving the global model on the server, the server broadcasts the new ω_S model to all clients in order to update the old client models. The whole learning process ends when the desired accuracy is reached.

B. Types

Based on our review of recent advances in FL in the energy domain, we classify its types from the data partitioning and network structure perspectives.

1) Data partition: Depending on how the training data is distributed in the sample space and feature space, FL can be classified into three categories: horizontal FL, vertical FL and federated transfer learning.

- **Horizontal FL**: In horizontal FL, the training data of the clients have the same feature space but different sample spaces. Due to this settings, clients can use the same machine learning model for jointly training the global FL

model. To ensure security, the local gradient of the clients usually needs to be protected by encryption techniques [9].

- **Vertical FL**: Vertical FL emphasizes shared global model learning in a network of clients with the same sample space and different data feature spaces [17]. In vertical FL, a specific alignment method should be designed to collect overlapping data samples from clients. These samples are combined to train a general machine learning model using encryption techniques.

- **Federated transfer learning**: Federated transfer learning is designed with the aim of extending the sample space in the vertical FL architecture to more learning clients for data with different sample spaces and different feature spaces [18]. In the concept of federated transfer learning, different feature spaces should be transformed into the same representation to aggregate data from multiple clients. To protect data privacy, encryption techniques are also usually widely used.

2) Network structure: Depending on the network topology, FL can be divided into two categories: central server-based FL and distributed FL.

- **Central server-based FL**: Central server-based FL is one of the most popular FL architectures. Fig. 1 is a typical structure of the central server-based FL. In this type of FL, there is a central server and a set of clients that execute the FL model. During a training round, all clients participate in the training of the network model in parallel using their own data. All clients then transmit the trained parameters to the central server, which aggregates these parameters using a weighted average algorithm such as FedAvg. Then, the computed global model is sent back to all clients for the next training iteration. At the end of the training process, each client receives an identical global model and its personalized model. This type of FL is highly dependent on the server model to coordinate the aggregation and distribute model updates.

- **Distributed FL**: Distributed FL is a network topology without any central server to coordinate the training process. Instead, all clients are connected in a peer-to-peer (P2P) manner to perform AI training [19]. In this way, in each round of communication, clients are also trained locally on their own data. Each client then implements model aggregation using model updates received from neighboring clients via P2P communication to agree on global updates. A distributed FL is designed to completely or partially replace central server-based FL when communication with the server is unavailable or when the network topology is highly scalable. Due to its contemporary features, a distributed FL can be integrated with P2P-based communication technologies such as blockchain to build a decentralized FL system. In this way, distributed FL clients can communicate through a blockchain ledger where model updates can be offloaded to the blockchain for secure model exchange and aggregation.

C. Aggregation Algorithms

Aggregation algorithm plays a key role in FL [8]. It incorporates the updates of all local models from all participating clients. The aim of the aggregation algorithm is to: 1) enhance privacy from local model updates; 2) preserve communication bandwidth; 3) facilitate asynchronous updates from clients. As the core of FL, the implementation of the aggregation algorithm varies according to the pre-defined needs. The current most widely used aggregation algorithms can be summarized as follows.

- **FedAvg**: This algorithm was originally proposed by Google. In FedAvg, the central server as the coordinator initiates the FL training process by sharing global parameters or global models with each selected client. Each selected client uses local data to train a global model and shares the weights of the trained model with the central server. The update of the global model is calculated by averaging the local models collected on the central server [20].

- **FedProx**: This algorithm is an improved version of FedAvg, which is proposed to address heterogeneity in FL [21]. FedProx takes into account variations in computing power and different factors in devices participating in FL training rounds. FedProx also introduces a proximal term to handle inconsistencies in local updates. The results of the experiment [21] indicate that FedProx can achieve positive results in heterogeneous settings.

- **FedMA**: This algorithm is used to build a shared model to aggregate CNN and LSTM models in FL [22]. FedMA averages the model on the central server through hierarchical matching and averaging hidden elements, such as neurons and channels in the neural network. As shown in the experimental results, FedMA works well on heterogeneous clients, surpassing FedAvg and FedProx in several rounds of training.

III. USE CASES

Electrification plays a key role in the whole energy landscape as the most efficient and environment-friendly final energy form with higher controllability due to advances in many technologies, such as materials, power electronics, information and communication technology (ICT), etc. In this section, we will focus on FL applications in energy (most of which are in the smart grid). As a particular data-driven model, technology innovation will only bring value when it fits into a particular use case. Therefore, we will first briefly introduce the general conditions of the smart grid domain, which is under dramatic changes. The main players that shape the landscape of the smart grid will be identified. Next, emerging FL applications based on the current literature review will be summarized from a use case perspective, along with their particular technology contributions. Special attention is paid to the aggregation mechanism for each FL application, as different use cases usually need customized FL frameworks with specific aggregation models.

A. *Smart Grid Landscape in the IoE and Its Stakeholders*

The power system witnesses radical revolutions featured by the "3D"–decolonization, decentralization, digitization, and it renders the system into a smart grid. Traditional stakeholders
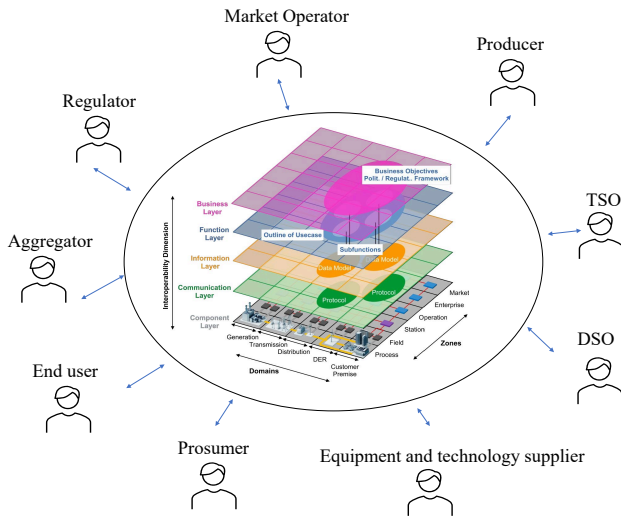
Fig. 2. Stakeholders in the Smart Grid

in the power sector are also changing due to the structural change from a vertical transmission of the energy starting from large power plants to end users through transmission lines, to a system which is networked with distributed energy resources. At the same time, more and more new players are entering the energy sector with an increasing deregulated and competitive electricity market. Among these new players, the interests of the prosumers who both consume and generate energy with distributed generations are becoming increasingly non-trivial.

As electricity continues to replace other energy forms as the final energy usage type, the interaction between the power sector and other energy sectors becomes more significant. It is no exaggeration to say that the smart grid will be the central enabler in the ecosystem of IoE, with many emerging technologies and applications, such as microgrids as the energy hubs, data analysis, multi-energy flow energy management, etc [23]–[25]. Fig. 2 depicts the current smart grid landscape that engages with different stakeholders in dimensions including interoperability, domains, and zones.

B. Use Case Summary

FL is an emerging research field in the energy sector. From a data analysis point of view, applications can be divided into analysis, prediction, and decision making (control and optimization). In this paper, we review the FL use cases for energy systems and summarize in Table I according to the use scenarios, the stakeholders benefited, challenges, FL types, and the aggregation feature of FL.

IV. CHALLENGES AND OPPORTUNITIES

FL is a new paradigm of distributed learning across different devices, data centers, and geographic locations. It is also a multidisciplinary learning framework that involves several disciplines, including cryptography, databases, and machine learning. Among many others, the challenges and opportunities of FL are summarized as follows: 1) How to improve the efficiency of the communication between the central server and the clients, which mainly involves how to reduce the

number of communications and compress the data. 2) How to protect the security of the data on the communication during the model training process. In this regard, data encryption techniques can be applied, including homomorphic encryption and differential privacy. 3) How to partition the data on the client side to achieve optimal parallel learning efficiency. Existing studies focus on the application of horizontal and vertical data partitioning strategies. 4) How to deal with non-identical distribution problems of the training data. These can be problems of skewed feature distribution and/or skewed label distribution. 5) How to improve learning efficiency, which mainly consists of studying how to optimize the machine learning algorithm and accelerate model convergence. 6) How to achieve multitask learning and personalized learning in a federated setting. Since clients may have different models, settings, and devices, this heterogeneity determines the challenges of balancing parallel tasks, thus providing personalized learning opportunities. 7) How to design machine learning algorithms and tune hyperparameters in an FL framework.

The above are the main challenges and opportunities for FL technology itself, which has been the subject of many research efforts. As a result, this has led to a large number of review articles; in this paper, we instead have focused on the other aspect that we think is most interesting, but has not yet been examined, namely, the global model generation algorithms. Aggregation algorithms are directly related to the performance of the final model, such as generality. In addition, we believe that FL should be fully decentralized or support peer-to-peer distributed learning, similar to blockchain technology where learning is consensus-based. Finally, due to the complexity of implementing FL technology and the applications it supports, we believe it would be beneficial to integrate visualization into an FL framework, which can greatly lower the learning curve and improve the interpretability of machine learning models implemented in an FL framework.

FL has a wide range of potential in the field of energy, which can be summarized as follows. First, it is often difficult to create powerful data-driven models for renewable energy prediction, mainly because of inherent stochasticity. For example, wind power generation is affected by wind speed, and solar photovoltaic energy is affected by the intensity and timing of solar exposure. Training a forecasting model with high accuracy and generality requires different data sources, which may be from power plants at different geographical locations. However, these data may not be available due to data ownership or commercial competition, in which case FL becomes essential to create a collaborative prediction model. Second, in the residential building sector, FL can help protect the privacy of customer information. For example, in Europe, the residential electricity supply is often operated by different distributors, each of which has its own customers. In order to create data-driven models from energy consumption and customer information, FL can address the model in a non-sharing environment with data privacy guarantee. This makes it possible to provide personalized energy services, such as energy-efficiency product recommendations, demand-

TABLE I
SUMMARY OF FEDERATED LEARNING FOR ENERGY SYSTEMS

Ref.	Scenarios	Stakeholder	Challenges	FL type	Aggregation feature
[26]	Household characteristic identification, such as income level, flat type, floor area, number and type of appliance and number of residence	Retailers	Data ownership, privacy preservation	Horizontal, central server-based FL	Asynchronous stochastic gradient descent with delay compensation (ASGD-DC)
[27]	Energy management of multiple smart homes equipped with DER in a system with one home energy management and several local home energy systems.	Aggregator of LHEMSs	computation burden, timeliness/real time, privacy preserve	FRL, horizontal	FedSGD algorithm
[28]	Residential building load forecasting with data from IoT devices	End user(smart home owner)	data access, privacy preserve, cloud service cost	Horizontal, decentralized FL	Decentralized, global server free
[29]	Load forecasting using smart meter data	utility	computation burden, data ownership, privacy preserve	horizontal, Central server based FL	FedAVG
[30]	Market participate and energy exchange with prosumer community group	Aggregator-prosumer community groups, end user	privacy preserve, data access	Horizontal, central server based FL	Averaging
[31]	Identificaiton of socio-demographic characteristics of electricity consumers from smart meter data	load forecasting using smart meter data	privacy preserve	Horizontal, central server based FL	Model averaging
[32]	Demand response through load control of multiple households to reduce peak load and user cost	Aggregator, end user(residential)	privacy preserve	Horizontal, central server based FL, RL	decentralized algorithm based on proximal Jacobian alternating direction method of multipliers (PJ-ADMM)
[33]	Behind-the-meter solar disaggregation from the net load to the solar generation profile at the community level	Utility	Privacy preserve	Horizontal, central server based FL	A layerwise parameter aggregation strategy
[27]	Home Energy management System(HEMS) of multiple smart homes equipped with DER.	Aggregato (HEMS that consists of a single GS and N LHEMSs)	computation burden, real time, privacy preserve	FRL, horizontal	FedSGD algorithm
[2]	Blade icing detection for wind turbines in the wind farms	wind farm owners	data ownership, computation burden	Horizontal, central server based FL	Differentiated computational capabilities in server and client models
[34]	Spatiotemporal scenario generation of renewable energies	Utility	privacy preserve	Horizontal, central server based FL	Integrating federated learning and least square generative adversarial networks
[35]	Probabilistic prediction of energy demand and driving range for electric vehicles to deal with limited driving range in a sparse charging infrastructure	EV owner	privacy preserve, computation and communication overhead	Horizontal, central server based FL	Federated Learning With Clustering, FedAvg and FedAG
[36]	Energy theft detection in smart grid	Utility	privacy preserve	Horizontal, central server based FL	FL with homomorphic encryption
[37]	Residential short-term energy consumption forecasting	Utility	privacy preserve	Horizontal, central server based FL	FedAVG
[38]	Non-intrusive load monitoring	Utility	data access, privacy preserve	Horizontal, central server based FL	FedAVG
[39]	Predicting energy demand for electric vehicle (EV) networks	Grid operator	privacy preserve	Horizontal, central server based FL	Clustering-based FL
[40]	Anomaly Detection in Smart Buildings	End user-the owner of the smart buildings	computation burden, timeliness,	Horizontal, central server based FL	FL with encryption(PySyft)
[41]	False data injection attack detection on power grid state estimation	Utility-power system state owner, detection service provider (DSP)	data access, privacy preserve, data ownership	Horizontal, central server based FL	Federated learning with incentive

response programs, and customized energy management solutions. Although research has been conducted on personalized recommendations based on FL in other fields, its applications in the energy sector remain an open problem that is interesting to investigate. Finally, energy involves many distributed optimization problems, including production, transportation, and distribution. The advantage of applying federated-based optimization is that it supports heterogeneous models, i.e., the model structure does not need to be consistent across different nodes. Therefore, an ensemble-like heterogeneous distributed optimization model can be established to help with decision making regarding energy planning or supply.

We are aware that energy is a very broad concept that has many subdomains. For example, depending on the types of energy, it can be divided into water, electricity, gas, heat, etc., each with its own life cycle for an energy system. In addition, there are many other energy-related research areas, including market, price, regulation, and policies. Therefore, FL has a wide potential in energy.

V. CONCLUSION AND FUTURE WORK

FL is an emerging computing paradigm. In this paper, we presented a systematic review of FL and its applications in the energy sector. We first introduced the concept of FL, the technical advances especially about the research of solutions, network structures, and global model generation algorithms. Then, we investigated various solutions and applications of FL in energy, mainly involving demand-response programs, user characteristics identification, customer segmentation, energy prediction, and optimization. We also discussed the technical challenges and opportunities of the FL technology itself and enumerated its opportunities in the energy sector. We hope that this survey will help researchers refine FL techniques and provide useful reference material for researchers in the energy field.

In the future work, we will improve our review in a more systematic way and update the latest research advances on FL and its applications in energy. We also intend to develop an FL benchmarking and evaluation tool, as no such tool is available to date.

REFERENCES

[1] Y. Li, S. Kubicki, A. Guerriero, and Y. Rezgui, "Review of building energy performance certification schemes towards future improvement," *Renewable and Sustainable Energy Reviews*, vol. 113, p. 109244, 2019.

[2] X. Cheng, F. Shi, Y. Liu, J. Zhou, X. Liu, and L. Huang, "A class-imbalanced heterogeneous federated learning model for detecting icing on wind turbine blades," *IEEE Transactions on Industrial Informatics*, 2022.

[3] P. Voigt and A. Von dem Bussche, "The eu general data protection regulation (gdpr)," *A Practical Guide, 1st Ed., Cham: Springer International Publishing*, vol. 10, no. 3152676, pp. 10–5555, 2017.

[4] J. Konečnỳ, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," *arXiv preprint arXiv:1610.05492*, 2016.

[5] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 50–60, 2020.

[6] X. Yin, Y. Zhu, and J. Hu, "A comprehensive survey of privacy-preserving federated learning: A taxonomy, review, and future directions," *ACM Computing Surveys (CSUR)*, vol. 54, no. 6, pp. 1–36, 2021.

[7] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, and Y. Gao, "A survey on federated learning," *Knowledge-Based Systems*, vol. 216, p. 106775, 2021.

[8] V. Mothukuri, R. M. Parizi, S. Pouriyeh, Y. Huang, A. Dehghantanha, and G. Srivastava, "A survey on security and privacy of federated learning," *Future Generation Computer Systems*, vol. 115, pp. 619–640, 2021.

[9] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, and H. V. Poor, "Federated learning for internet of things: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, 2021.

[10] K. J. Rahman, F. Ahmed, N. Akhter, M. Hasan, R. Amin, K. E. Aziz, A. M. Islam, M. S. H. Mukta, and A. N. Islam, "Challenges, applications and design aspects of federated learning: A survey," *IEEE Access*, vol. 9, pp. 124 682–124 700, 2021.

[11] L. Li, Y. Fan, M. Tse, and K.-Y. Lin, "A review of applications in federated learning," *Computers & Industrial Engineering*, vol. 149, p. 106854, 2020.

[12] F. A. KhoKhar, J. H. Shah, M. A. Khan, M. Sharif, U. Tariq, and S. Kadry, "A review on federated learning towards image processing," *Computers & Electrical Engineering*, vol. 99, p. 107818, 2022.

[13] R. S. Antunes, C. A. da Costa, A. Küderle, I. A. Yari, and B. Eskofier, "Federated learning for healthcare: Systematic review and architecture proposal," *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2022.

[14] M. Ali, F. Naeem, M. Tariq, and G. Kaddoum, "Federated learning for privacy preservation in smart healthcare systems: A comprehensive survey," *arXiv preprint arXiv:2203.09702*, 2022.

[15] M. Liu, S. Ho, M. Wang, L. Gao, Y. Jin, and H. Zhang, "Federated learning meets natural language processing: A survey," *arXiv preprint arXiv:2107.12603*, 2021.

[16] X. Cheng, F. Shi, Y. Liu, X. Liu, and L. Huang, "Wind turbine blade icing detection: a federated learning approach," *Energy*, vol. 254, p. 124441, 2022.

[17] S. Feng and H. Yu, "Multi-participant multi-class vertical federated learning," *arXiv preprint arXiv:2001.11154*, 2020.

[18] Y. Liu, Y. Kang, C. Xing, T. Chen, and Q. Yang, "A secure federated transfer learning framework," *IEEE Intelligent Systems*, vol. 35, no. 4, pp. 70–82, 2020.

[19] X. Cheng, W. Tian, F. Shi, M. Zhao, S. Chen, and H. Wang, "A blockchain-empowered cluster-based federated learning model for blade icing estimation on iot-enabled wind turbine," *IEEE Transactions on Industrial Informatics*, 2022.

[20] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*. PMLR, 2017, pp. 1273–1282.

[21] T. Li, A. K. Sahu, M. Zaheer, M. Sanjabi, A. Talwalkar, and V. Smith, "Federated optimization in heterogeneous networks," *Proceedings of Machine Learning and Systems*, vol. 2, pp. 429–450, 2020.

[22] H. Wang, M. Yurochkin, Y. Sun, D. Papailiopoulos, and Y. Khazaeni, "Federated learning with matched averaging," *arXiv preprint arXiv:2002.06440*, 2020.

[23] C. Li, O. B. Fosso, M. Molinas, Y. Jingpeng, and P. Raboni, "Defining three distribution system scenarios for microgrid applications," in *2020 IEEE 4th Conference on Energy Internet and Energy System Integration (EI2)*. IEEE, 2020, pp. 982–987.

[24] C. Li, M. Molinas, O. B. Fosso, N. Qin, and L. Zhu, "A data-driven approach to grid impedance identification for impedance-based stability analysis under different frequency ranges," in *2019 IEEE Milan PowerTech*. IEEE, 2019, pp. 1–6.

[25] C. Li, S. K. Chaudhary, M. Savaghebi, J. C. Vasquez, and J. M. Guerrero, "Power flow analysis for low-voltage ac and dc microgrids considering droop control and virtual impedance," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2754–2764, 2016.

[26] J. Lin, J. Ma, and J. Zhu, "Privacy-Preserving Household Characteristic Identification With Federated Learning Method," *IEEE Transactions on Smart Grid*, vol. 13, no. 2, pp. 1088–1099, 2022.

[27] S. Lee and D.-H. Choi, "Federated Reinforcement Learning for Energy Management of Multiple Smart Homes With Distributed Energy Resources," *IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS*, vol. 18, no. 1, pp. 488–497, Jan. 2022.

[28] J. Gao, J. Gao, W. Wang, Z. Liu, F. R. M. Billah, and B. Campbell, "Decentralized Federated Learning Framework for the Neighborhood: A Case Study on Residential Building Load Forecasting," *SenSys*, 2021.

[29] M. N. Fekri, K. Grolinger, and S. Mir, "Distributed load forecasting using smart meter data: Federated learning with Recurrent Neural Networks," *International Journal of Electrical Power & Energy Systems*, 2021.

[30] A. Taik, B. Nour, and S. Cherkaoui, "EMPOWERING PROSUMER COMMUNITIES IN SMART GRID WITH WIRELESS COMMUNICATIONS AND FEDERATED EDGE LEARNING," *IEEE WIRELESS COMMUNICATIONS*, vol. 28, no. 6, pp. 26–33, Dec. 2021.

[31] Y. Wang, I. L. Bennani, X. Liu, M. Sun, and Y. Zhou, "Electricity Consumer Characteristics Identification: A Federated Learning Approach," *IEEE TRANSACTIONS ON SMART GRID*, vol. 12, no. 4, pp. 3637–3647, Jul. 2021.

[32] S. Bahrami, Y. C. Chen, and V. W. S. Wong, "Deep Reinforcement Learning for Demand Response in Distribution Networks," *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1496–1506, 2021.

[33] J. Lin, J. Ma, and J. Zhu, "A Privacy-Preserving Federated Learning Method for Probabilistic Community-Level Behind-the-Meter Solar Generation Disaggregation," *IEEE Transactions on Smart Grid*, vol. 13, no. 1, pp. 268–279, 2022.

[34] Y. Li, Y. Li, J. Li, J. Li, and Y. Wang, "Privacy-preserving Spatiotemporal Scenario Generation of Renewable Energies: A Federated Deep Generative Learning Approach." *IEEE Transactions on Industrial Informatics*, 2021.

[35] A. T. Thorgeirsson, A. T. Thorgeirsson, A. T. Thorgeirsson, S. Scheubner, S. Funfgeld, and F. Gauterin, "Probabilistic Prediction of Energy Demand and Driving Range for Electric Vehicles With Federated Learning," *IEEE Open Journal of Vehicular Technology*, 2021.

[36] M. Wen, R. Xie, K. Lu, K. Lu, L. Wang, L. Wang, K. Zhang, and K. Zhang, "FedDetect: A Novel Privacy-Preserving Federated Learning Framework for Energy Theft Detection in Smart Grid," *IEEE Internet of Things Journal*, 2021.

[37] M. Savi and F. Olivadese, "Short-Term Energy Consumption Forecasting at the Edge: A Federated Learning Approach," *IEEE access : practical innovations, open solutions*, vol. 9, pp. 95 949–95 969, 2021.

[38] H. Wang, C. Si, and J. Zhao, "Fed-NILM: A Federated Learning-based Non-Intrusive Load Monitoring Method for Privacy-Protection," *Energy Conversion and Economics*, 2021.

[39] Y. M. Saputra, D. T. Hoang, D. N. Nguyen, E. Dutkiewicz, M. D. Mueck, M. D. Mueck, and S. Srikanteswara, "Energy Demand Prediction with Federated Learning for Electric Vehicle Networks," *arXiv: Signal Processing*, 2019.

[40] R. A. Sater and A. B. Hamza, "A Federated Learning Approach to Anomaly Detection in Smart Buildings," *arXiv: Learning*, 2020.

[41] W.-T. Lin, W.-T. Lin, W.-T. Lin, G. Chen, and Y. Huang, "Incentive edge-based federated learning for false data injection attack detection on power grid state estimation: A novel mechanism design approach," *Applied Energy*, 2022.