



Why Reinforcement Learning in Energy Systems Needs Explanations

Hallah Shahid Butt
hallah.butt@kit.edu

Karlsruhe Institute of Technology
Germany

Benjamin Schäfer
benjamin.schaefer@kit.edu

Karlsruhe Institute of Technology
Germany

ABSTRACT

With economic development, the complexity of infrastructure has increased drastically. Similarly, with the shift from fossil fuels to renewable sources of energy, there is a dire need for such systems that not only predict and forecast with accuracy but also help in understanding the process of predictions. Artificial intelligence and machine learning techniques have helped in finding out well-performing solutions to different problems in the energy sector. However, the usage of state-of-the-art techniques like reinforcement learning is not surprisingly convincing. This paper discusses the application of reinforcement techniques in energy systems and how explanations of these models can be helpful.

KEYWORDS

explainable artificial intelligence (XAI), reinforcement learning, machine learning, interpretations, energy systems

ACM Reference Format:

Hallah Shahid Butt and Benjamin Schäfer. 2024. Why Reinforcement Learning in Energy Systems Needs Explanations. In *2024 Workshop on Explainability Engineering (ExEn '24)*, April 20, 2024, Lisbon, Portugal. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3648505.3648510>

1 INTRODUCTION

Infrastructure, such as hospitals, power plants, and universities, play an important role in our daily lives. Economic and global development has increased infrastructure complexity. With the help of advancements in technology and automation of labour-intensive tasks, these complex systems require electrical power to be functional and operative. Therefore, to meet the increasing demand for electrical power, policymakers are encouraged to think about other unconventional sources of energy that not only fulfill the demand but also are more environmentally friendly. Conventional sources of energy, such as oil, gas, and coal energy, are not good for the environment as they emit immense amounts of CO₂. Moreover, the reserves of conventional resources are depleting with time, and this scarcity of resources can lead to severe problems for future generations. Consequently, sustainable distributed energy systems (DERs) are integrated into the energy systems [1]. While

these systems emit less CO₂, their inclusion has significantly increased the complexity of the power grids since these DERs are distributed across the power system (potentially far away from consumers) and are weather-dependent. This later property of the DERs adds uncertainty to their nature; hence, it demands the need for proper management and planning of the systems to meet the power demand and supply.

In 1956 when the term "Artificial intelligence (AI)" was coined [2], it was presented as an idea of simulating learning and solving every real-world problem on computers [3] with less or no human intervention [2]. Since computational resources have become readily available, it has become easy for computers to learn the problem-solving process and solve complex tasks in diverse areas with the help of AI and machine learning (ML) algorithms. These techniques have made a foray into every scientific field whether it is healthcare, banking [3], transportation [4], and many other areas. For instance, in the medical field, researchers have applied AI and ML techniques to diagnose diseases like cancer [5], outbreak prediction [6], and medical image processing with the help of computer vision [7, 8]. Similarly, researchers have found opportunities to apply AI and ML techniques to address issues of "systems of systems".

AI has gained significant importance in scientifically addressing infrastructure issues. A lot of studies have been carried out on infrastructure systems and researchers have done an extensive review on how AI can assist us in dealing with problems arising in infrastructure systems like energy, transportation, water systems, and telecommunications [9]. In energy systems, AI and ML techniques can be used to tackle numerous challenges of the ongoing phase-out of conventional resource usage and the increasing usage of renewable energy. This includes predicting the power generation and load, enhancing monitoring of the system, and finding optimal solutions to reduce power losses [10]. The above-mentioned problems can be optimally solved with the help of state-of-the-art methods like reinforcement learning. RL methods have proven their credibility in other fields but in energy systems, their application is not convincing which we will discuss in a later section.

Despite the accuracy and precision of the models generated by advanced and state-of-the-art algorithms, policymakers and companies are unable to trust the outputs of many of these algorithms because of their black-box nature [11]. Even the developers of these black-box models are often unable to comprehend how the model processed the different input variables to convert to a certain output function and how the inputs are related to each other [12]. Interpretation of these models is crucial in all domains, especially in medicine, energy systems, and other complex infrastructures, as



This work licensed under Creative Commons Attribution International 4.0 License.

ExEn '24, April 20, 2024, Lisbon, Portugal
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0596-0/24/04.
<https://doi.org/10.1145/3648505.3648510>

we need to trust the systems concerned with such a vital part of our lives.

In this paper, we will discuss the applications of ML and RL in energy systems, explain why the interpretation of ML models is important, which techniques are used for explaining these models, and in how existing knowledge should be advanced by new research.

2 ENERGY SYSTEMS AND REINFORCEMENT LEARNING

As discussed earlier, energy systems have reached a level of complexity, and with more sensors, an increasing amount of data becomes available so that we can find solutions to these challenges to the energy system with the help of ML-based algorithms. Already, with the digitization of the systems, the usage of machine learning methods has increased drastically. Including DERs made it crucial to monitor the energy systems so they can operate effectively. To model these systems, it is necessary to have a good knowledge of the parameters on which such systems depend, necessitating inference [4].

The most commonly used learning types in energy systems are supervised, and unsupervised learning. These techniques are used to forecast the power generation, optimally manage the operational and maintenance cost [13], and schedule the storage and generation operations. Authors in [13] have reviewed the number of articles published to apply machine learning methods in the energy domain, and they have found out that after 2012, researchers are more interested in using ML techniques in energy systems, and the number of publications made annually reached up to 2000 [14]. Authors in [9] support this increased interest as they reviewed the work of researchers who have used ML-based methods to effectively forecast the generation, manage demand and supply, and monitor the performance of the systems. Hence, both reviews stress the importance of how AI and ML approaches have efficiently solved complex problems in energy systems.

Even though algorithms behave differently in different domains, authors in [9] have curated a list of well-performing algorithms that include artificial neural networks (ANN), support vector machines (SVM), decision trees and random forests, and K-mean clustering. Previously, in energy systems, the community leaned mostly toward traditional approaches. However, with the transition and advancements in the methods, researchers in the energy sector are also increasingly adopting ANN-based techniques. Areas like demand forecasting, weather predictions, and price forecasting are popular and are solved by these algorithms [9]. For example, Hafeez et al [15] have proposed a novel electrical load forecasting hybrid model that outperforms previous ANN and LSTM algorithms concerning the accuracy, runtime, and convergence rate. Furthermore, researchers have also reviewed and studied how ANN can be applied to optimize the performance of solar collectors [16]. They compared the performance with traditional techniques like linear regression and non-linear models and found that ANNs outperformed the other methods.

In addition to the use of supervised and unsupervised learning methods, RL is also attracting attention, and researchers are keen to implement RL methods in different domains. Though we have seen

RL-based methods applied in different sectors, the researchers of the energy domain seem to be reluctant to adopt RL approaches as can be seen from a low number of RL publications in energy systems [14]. In RL, the learning agent learns a specific policy based on the feedback given by the environment [17]. The RL agent receives a state from state space, and it takes actions either to explore the state space further or to exploit learned rules. This new action leads the agent to a new state after receiving a reward or penalty from the environment [18]. This data-driven method helps to learn and create new policies based on trial and error to attain the maximum reward. RL approaches are more suitable to find optimal solutions and react to dynamic environments than supervised or unsupervised learning. Moreover, in real-time decision-making and optimization problems, the RL-based method can be the best choice [19]. In recent times, RL methods have proven their worth in different fields like gaming, robotics [19, 20] traffic control systems, advertising, and the stock industry [19].

Since the complexity of energy systems escalated due to the inclusion of DERs, it is of key importance to optimally control these systems. Fortunately, the rise of complexity is accompanied by an ever growing amount of operational and real-time data. Processing and harnessing this bulk of data requires state-of-the-art ML and specifically RL approaches to control and optimize the systems [21]. From dispatching generated power optimally over the network to storing energy for later consumption, highly renewable energy systems are more complicated in terms of their system management and therefore offer numerous opportunities for RL applications. The uncertainties of the system like dependency on weather, energy market variations, and improved technologies of the energy systems, state-of-the-art techniques are required to address the problems [14]. Regardless of the applications and domain, reinforcement learning methodologies have attained success in the scientific field as these methodologies handle uncertainties. Furthermore, for some complex problems, model-free solutions are suitable rather than model-based approaches. Thus, RL methods are a powerful and effective technique for the modern challenges of energy systems [14].

Concluding, more data becomes available for energy systems, while the integration of DERs increases complexity and requires optimization and control solutions capable of processing real-time multi-dimensional data efficiently. RL-based approaches are uniquely suitable to address this challenge. The scalability of RL-based approaches has improved with the improvement of state-of-the-art methods like Deep Q network and Proximal Policy Optimization (PPO) since they deal with multi-dimensional data [20]. Consequently, we are also proposing to use Deep Q network and PPO to address our research question, as discussed in section 4.

3 ENERGY SYSTEMS AND INTERPRETATION OF MODELS

Because of the advancements in the field of data science, many data-driven complex tasks are solved with high-level accuracy with the help of ML algorithms. However, most of these models and output functions are uninterpretable, and even the designers of those algorithms are unable to understand how these inputs are converted into the following output. In ML, most of the models are

black-box, i.e. these models are generated in an untransparent way that does not reveal how results are obtained, e.g. because a large amount of intrinsically simple units (neural layers) are combined into one extensive complex system. This is in contrast to transparent (or white box) models, such as linear regression, where we clearly see how the model operates but which unfortunately are often performing worse than black boxes [12]. AI and ML techniques face backlash from being unreliable and vulnerable to bias. This behaviour of producing unfair and biased results is analyzed by studying AI-designed tools where a famous criminal risk assessment tool was biased in judging African-Americans as criminal, and a beauty pageant winners tool judging black-skinned toned participants unfairly [22, 23]. Such unreliable results can impact the decision-making process and it will lead to wrong decisions which can later on affect the process and management. Hence, it is of key importance to have interpretability and transparency of the models so that it can help in improving the credibility of models and elevate acceptance of their decisions.

In this digital age, it becomes inevitable to use AI-based systems for the decision-making processes. The intelligence of systems is helping decision-makers and policymakers to think in different ways. It is also evident that the fairness, transparency, and interpretability of the predictions or suggestions provided by the AI systems are crucial because these systems have a great impact on human lives [24]. According to [24], the most important domains where explainability of these systems are medical and law. However, this idea to make the models explainable is not confined to only these areas. As we discussed earlier, energy systems are also complex systems; hence, the application of AI and ML techniques is beneficial in the decision-making process. However, these ML-based systems are still not fully adopted in reality despite the accuracy and advancement in ML technologies because the black-box nature of such algorithms is a key hurdle [25]. Furthermore, the unreliability and lack of understanding of these systems is another barrier to adopting ML-based systems specifically in the energy domain [26]. So, to implement these ML-based systems, it is evident that there should be a system that is interpreted by humans and adds to model validation and inspection. Thus, increasing transparency of ML solutions will likely lead to the adaptation of more such models in the energy field.

The notion of explanations is defined differently, depending on the context. Sometimes it is presented as an explanation of an intrinsic process which helps users to understand the process [24], and some researchers categorize the explanations in various sub-categories to achieve different goals like generating trust and transparency, following some renowned regulations, ensuring fairness, generating accountable and reliable models, reduction of misinterpretation in model performance, and validating models [27]. In energy systems, we need explanations as just discussed (e.g. trust, reliability, fairness). Current methods do not yet provide full explanations but so far focus on deriving interpretations of the trained models that identify and quantify the importance of different features or input in generating any model as [24] has done work while modeling energy systems. The necessary explanations also vary depending on the target group, for instance, experts need explanations in terms of finding out the potential significant features in predictions as researchers in [28] have also interpreted their models

Years	XAI and Energy systems
2020-2023	64
2010-2019	4
2000-2009	0

Table 1: This table shows the number of publications published within these years that were related to keywords “XAI and energy systems“

by computing the feature importance, whereas, for a layperson, the underlying mechanics of the system are not important. Instead, they are interested in actionable and causal explanations of the predictions as mentioned in [26].

Though we are witnessing the increasing trend of applying explainable artificial intelligence (XAI) techniques in the energy field, it is still a new and emerging aspect in the field of energy systems, and a lot of work is yet to be done as it is shown in Table 1. It is important to note that these publications have used a variety of explainability approaches to elaborate a range of energy system problems using again a spectrum of different ML-based methods, including for example SHAP on regression or attention for forecasts [29]

More and more researchers are shifting their attention to explaining the ML models used in energy systems. There are some areas where the concept of interpreting the models is widely used, including power grid applications, the energy sector, and energy management in buildings [29]. Authors in [29] have listed the commonly used XAI approaches in the energy domain including Shapley additive explanations (SHAP) [30–33], feature correlation [34, 35], and permutation feature importance [26, 36].

Similar to the explanations of other ML techniques, reinforcement learning is also gaining importance in energy systems so that energy experts can build trust in the systems. Many researchers have implemented deep reinforcement learning (DRL) to solve complex problems. For example, Huang et al [20] have addressed the power energy control system with the help of DRL methods. While in [37], researchers have worked on implementing DRL to address power emergency control, and they have used deep explainer to obtain at least initial interpretations of previously black box models. However, the explainability of such systems is still mostly missing. Thus, a lot of work has yet to be done since interpretations, let alone full explanations, of reinforcement learning algorithms in energy are not very common, as suitable methods are not readily available. To name two examples of RL applications in energy systems that require explanations:

- First, an RL controller of an energy storage system needs explanations because this will help energy experts or laypeople understand and trust the control laws which are learnt by the system.
- Second, frequency and voltage control of transmission systems could be realized by an RL agent, especially in emergency situations, where humans are too slow to react and existing rule-based systems are too inflexible.

However, actions to stabilize a system have to convince operators so that they take appropriate action and so that potential damage due to failure can be properly investigated. Both of the above-mentioned

RL applications make explanations for these models inevitable [29].

4 OUR PLANNED CONTRIBUTION

We wish to focus on explainable reinforcement learning in small-scale energy systems. Due to the stochastic nature of renewable energy sources, it is natural to include energy storage in the energy system to ensure a smoother and more constant supply of power, e.g. to a household. Thermal energy storage (TES) systems are one candidate for balancing the supply and demand of the buildings. Wang et al have already developed a deep Q-network RL approach to optimally control this setup [38]. Utilizing another technology, Elseify et al [39] considered a scenario with multiple batteries to store photovoltaic (PV) generated power and again obtained an optimal solution using RL methods, specifically an equilibrium optimizer. However, it is crucial to derive explanations for experts and laypeople alike. Hence, we require suitable methods and proofs-of-concept for explainable reinforcement learning [29].

In our research, we want to identify the optimal power management of a building, including balancing renewable generation and consumption via demand response and storage. Currently, we are considering the use-case of buildings with roof-mounted PV panels that have to balance PV generation, storage and purchases from the power grid. We further assume that these buildings have installed batteries to store excess solar power generated by the roof-mounted PV so that this power can later be used when insufficient power is generated by the PV panels, especially at night and on cloudy days. We search for optimal usage strategies of batteries to minimize the total costs of purchasing power from the connected utility grid. Rather than using supervised and rule-based approaches, we aim to solve this optimization problem with the help of state-of-the-art RL methods like Deep Q-network, and PPO [40], so that the agent learns the policies as per the given states of the environment. Technically, we are modeling a single building in terms of demand D , grid electricity prices p , stored energy in the batteries, and the power generated by the PV panel G_{PV} . We have used the notion of net demand which is the function of actual demand and power generation of PV panel.

$$\text{Net demand} = f(D, G_{PV}) = D - G_{PV}. \quad (1)$$

This net demand indicates the extra power that the building needs from the grid. The agent takes specific actions, for instance, to charge and discharge the battery, and these actions are either rewarded or penalized by the environment.

However, we also want to derive useful interpretations and explanations from the model which is a black-box model. Initial steps include interpreting the workings of the model, e.g. by obtaining a ranking of features that have an impact on the predictions with the help of break-down profiles [41]. Thus, we can formulate our research question as

Can we use RL to solve an optimization problem in a small-scale energy system (balancing power of a single building) and how can we improve trust in such a black-box model by explaining its learned policies and actions?

In summary, our research contributes in the following way:

- Analysing the applicability of RL in solving optimization problems in energy systems, like minimizing the cost of purchasing electricity from the grid by maximizing the consumption of PV panels' generated power
- Having a comparative analysis of different state-of-the-art RL methods
- Having explanations of the black-box models to build trust in the outcomes of these models and analysing how features impact the predictions

Later on, we will look for full explanations of why the model took a certain action and why it did not take a different one (counterfactual explanations), and ideally, generate these explanations dynamically to explain the rationale of each step and every point in time. The goal is to increase trust and transparency so that these systems can be implemented at large scales and experts and laypeople feel confident while adapting such systems.

5 CONCLUSION

Energy systems, as critical infrastructure and with many intrinsic optimization problems, are excellent systems to apply explainable ML and RL techniques. Unfortunately, XAI is still used too reluctantly for supervised problems and even more rarely for RL. Hence, we propose to bridge this gap and develop suitable explainable reinforcement learning methods for energy systems.

6 ACKNOWLEDGEMENTS

We gratefully acknowledge funding from the Helmholtz Association under grant no. VH-NG-1727 and the Networking Fund through Helmholtz AI.

REFERENCES

- [1] H. Shengren, P. P. Vergara, E. M. S. Duque, and P. Palensky, "Optimal energy system scheduling using a constraint-aware reinforcement learning algorithm," *International Journal of Electrical Power & Energy Systems*, vol. 152, p. 109230, 2023.
- [2] P. Hamet and J. Tremblay, "Artificial intelligence in medicine," *Metabolism*, vol. 69, pp. S36–S40, 2017.
- [3] A. Entezari, A. Aslani, R. Zahedi, and Y. Noorollahi, "Artificial intelligence and machine learning in energy systems: A bibliographic perspective," *Energy Strategy Reviews*, vol. 45, p. 101017, 2023.
- [4] M. M. Forootan, I. Larki, R. Zahedi, and A. Ahmadi, "Machine learning and deep learning in energy systems: A review," *Sustainability*, vol. 14, no. 8, 2022. [Online]. Available: <https://www.mdpi.com/2071-1050/14/8/4832>
- [5] S. Huang, J. Yang, S. Fong, and Q. Zhao, "Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges," *Cancer letters*, vol. 471, pp. 61–71, 2020.
- [6] J. K. K. Nayak, A. K. Nayak, S. K. Nayak, V. Shaj, and S. N. Karthikeyan, "Epidemic outbreak prediction using artificial intelligence," *International Journal of Computer Science & Information Technology (IJCSIT)*, vol. 10, no. 4, pp. 49–64, 2018.
- [7] J. Olveres, G. González, F. Torres, J. C. Moreno-Tagle, E. Carbajal-Degante, A. Valencia-Rodríguez, N. Méndez-Sánchez, and B. Escalante-Ramírez, "What is new in computer vision and artificial intelligence in medical image analysis applications," *Quantitative imaging in medicine and surgery*, vol. 11, no. 8, p. 3830, 2021.
- [8] D. Sathya, V. Sudha, and D. Jagadeesan, "Application of machine learning techniques in healthcare," in *Handbook of Research on Applications and Implementations of Machine Learning Techniques*. IGI Global, 2020, pp. 289–304.
- [9] L. McMillan and L. Varga, "A review of the use of artificial intelligence methods in infrastructure systems," *Engineering Applications of Artificial Intelligence*, vol. 116, p. 105472, 2022.
- [10] D. Rolnick, P. L. Donti, L. H. Kaack, K. Kochanski, A. Lacoste, K. Sankaran, A. S. Ross, N. Milojevic-Dupont, N. Jaques, A. Waldman-Brown, A. S. Luccioni, T. Maharaj, E. D. Sherwin, S. K. Mukkavilli, K. P. Kording, C. P. Gomes, A. Y. Ng, D. Hassabis, J. C. Platt, F. Creutzig, J. Chayes, and Y. Bengio, "Tackling climate

- change with machine learning,” *ACM Comput. Surv.*, vol. 55, no. 2, feb 2022. [Online]. Available: <https://doi.org/10.1145/3485128>
- [11] R. Abduljabbar, H. Dia, S. Liyanage, and S. A. Bagloee, “Applications of artificial intelligence in transport: An overview,” *Sustainability*, vol. 11, no. 1, p. 189, 2019.
 - [12] C. Rudin and J. Radin, “Why are we using black box models in ai when we don’t need to? a lesson from an explainable ai competition,” *Harvard Data Science Review*, vol. 1, no. 2, pp. 1–9, 2019.
 - [13] W. Shin, J. Han, and W. Rhee, “Ai-assistance for predictive maintenance of renewable energy systems,” *Energy*, vol. 221, p. 119775, 2021.
 - [14] A. Perera and P. Kamalaruban, “Applications of reinforcement learning in energy systems,” *Renewable and Sustainable Energy Reviews*, vol. 137, p. 110618, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1364032120309023>
 - [15] G. Hafeez, K. S. Alimgeer, and I. Khan, “Electric load forecasting based on deep learning and optimized by heuristic algorithm in smart grid,” *Applied Energy*, vol. 269, p. 114915, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S030626192030427X>
 - [16] H. K. Ghrtilahre and R. K. Prasad, “Application of ann technique to predict the performance of solar collector systems - a review,” *Renewable and Sustainable Energy Reviews*, vol. 84, pp. 75–88, 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1364032118300017>
 - [17] L. P. Kaelbling, M. L. Littman, and A. W. Moore, “Reinforcement learning: A survey,” *Journal of artificial intelligence research*, vol. 4, pp. 237–285, 1996.
 - [18] Y. Li, “Deep reinforcement learning: An overview,” *arXiv preprint arXiv:1701.07274*, 2017.
 - [19] C. Dilmegani, “Reinforcement learning: Benefits & applications in 2024,” <https://research.aimultiple.com/reinforcement-learning/>, 2022.
 - [20] Q. Huang, R. Huang, W. Hao, J. Tan, R. Fan, and Z. Huang, “Adaptive power system emergency control using deep reinforcement learning,” *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1171–1182, 2020.
 - [21] D. Cao, W. Hu, J. Zhao, G. Zhang, B. Zhang, Z. Liu, Z. Chen, and F. Blaabjerg, “Reinforcement learning and its applications in modern power and energy systems: A review,” *Journal of modern power systems and clean energy*, vol. 8, no. 6, pp. 1029–1042, 2020.
 - [22] L. H. Gilpin, D. Bau, B. Z. Yuan, A. Bajwa, M. Specter, and L. Kagal, “Explaining explanations: An overview of interpretability of machine learning,” in *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, 2018, pp. 80–89.
 - [23] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan, “A survey on bias and fairness in machine learning,” *ACM Comput. Surv.*, vol. 54, no. 6, jul 2021. [Online]. Available: <https://doi.org/10.1145/3457607>
 - [24] J. J. Ferreira and M. Monteiro, “The human-ai relationship in decision-making: Ai explanation to support people on justifying their decisions,” *arXiv preprint arXiv:2102.05460*, 2021.
 - [25] W. Zhang, F. Liu, Y. Wen, and B. Nee, “Toward explainable and interpretable building energy modelling: An explainable artificial intelligence approach,” in *Proceedings of the 8th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, ser. BuildSys ’21. New York, NY, USA: Association for Computing Machinery, 2021, p. 255–258. [Online]. Available: <https://doi.org/10.1145/3488661.3491127>
 - [26] S. Wenninger, C. Kaymakci, and C. Wiethel, “Explainable long-term building energy consumption prediction using qlattice,” *Applied Energy*, vol. 308, p. 118300, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0306261921015592>
 - [27] J. Gerlings, A. Shollo, and I. Constantiou, “Reviewing the need for explainable artificial intelligence (xai),” *arXiv preprint arXiv:2012.01007*, 2020.
 - [28] S. Pütz, J. Kruse, D. Witthaut, V. Hagenmeyer, and B. Schäfer, “Regulatory changes in german and austrian power systems explored with explainable artificial intelligence,” in *Companion Proceedings of the 14th ACM International Conference on Future Energy Systems*, ser. e-Energy ’23 Companion. New York, NY, USA: Association for Computing Machinery, 2023, p. 26–31. [Online]. Available: <https://doi.org/10.1145/3599733.3600247>
 - [29] R. Machlev, L. Heistrene, M. Perl, K. Levy, J. Belikov, S. Mannor, and Y. Levron, “Explainable artificial intelligence (xai) techniques for energy and power systems: Review, challenges and opportunities,” *Energy and AI*, vol. 9, p. 100169, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2666546822000246>
 - [30] M. Kuzlu, U. Cali, V. Sharma, and O. Güler, “Gaining insight into solar photovoltaic power generation forecasting utilizing explainable artificial intelligence tools,” *IEEE Access*, vol. 8, pp. 187 814–187 823, 2020.
 - [31] G. Mitrentsis and H. Lens, “An interpretable probabilistic model for short-term solar power forecasting using natural gradient boosting,” *Applied Energy*, vol. 309, p. 118473, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0306261921016950>
 - [32] G. Alova, P. A. Trotter, and A. Money, “A machine-learning approach to predicting africa’s electricity mix based on planned power plants and their chances of success,” *Nature Energy*, vol. 6, no. 2, pp. 158–166, 2021.
 - [33] J. Kruse, S. Mandelli, and S. R. Khosravirad, “Qos-aware wireless sensor networks: Reliability and low-latency for heterogeneous industry 4.0,” in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, 2021, pp. 1–5.
 - [34] I. C. Konstantakopoulos, H. P. Das, A. R. Barkan, S. He, T. Veeravalli, H. Liu, A. B. Manasawala, Y.-W. Lin, and C. J. Spanos, “Design, benchmarking and explainability analysis of a game-theoretic framework towards energy efficiency in smart infrastructure,” *arXiv preprint arXiv:1910.07899*, 2019.
 - [35] H. P. Das, I. C. Konstantakopoulos, A. B. Manasawala, T. Veeravalli, H. Liu, and C. J. Spanos, “A novel graphical lasso based approach towards segmentation analysis in energy game-theoretic frameworks,” in *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*. IEEE, 2019, pp. 1702–1709.
 - [36] M. Chaibi, E. M. Benghoulam, L. Tarik, M. Berrada, and A. E. Hmaid, “An interpretable machine learning model for daily global solar radiation prediction,” *Energies*, vol. 14, no. 21, 2021. [Online]. Available: <https://www.mdpi.com/1996-1073/14/21/7367>
 - [37] K. Zhang, P. Xu, and J. Zhang, “Explainable ai in deep reinforcement learning models: A shap method applied in power system emergency control,” in *2020 IEEE 4th Conference on Energy Internet and Energy System Integration (EI2)*, 2020, pp. 711–716.
 - [38] X. Wang, X. Kang, J. An, H. Chen, and D. Yan, “Reinforcement learning approach for optimal control of ice-based thermal energy storage (tes) systems in commercial buildings,” *Energy and Buildings*, vol. 301, p. 113696, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S037877882300926X>
 - [39] M. A. Elseify, A. SeyyedAbbasi, J. L. Dominguez-García, and S. Kamel, “Probabilistic optimal planning of multiple photovoltaics and battery energy storage systems in distribution networks: A boosted equilibrium optimizer with time-variant load models,” *Journal of Energy Storage*, vol. 73, p. 108986, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352152X23023848>
 - [40] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” *arXiv preprint arXiv:1707.06347*, 2017.
 - [41] P. Biecek and T. Burzykowski, “6 break-down plots for additive attributions | explanatory model analysis,” <https://ema.drwhy.ai/breakDown.html>, 2020.