Les Cahiers du GERAD ISSN: 0711–2440

(Trustworthy) AI for Québec's virtual power plant

J. Pallage, A. Lesage-Landry

G–2024–23

March 2024

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2024
– Bibliothèque et Archives Canada, 2024

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2024
– Library and Archives Canada, 2024

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

(Trustworthy) AI for Québec's virtual power plant

Pallage, Julien

Lesage-Landry, Antoine

Department of Electrical Engineering, Polytechnique Montréal, Montréal (Qc), Canada, H3T 1J4

GERAD, Montréal (Qc), Canada, H3T 1J4

MILA, Montréal (Qc), Canada, H2S 3H1

julien.pallage@polymtl.ca
antoine.lesage-landry@polymtl.ca

March 2024
Les Cahiers du GERAD
G–2024–23

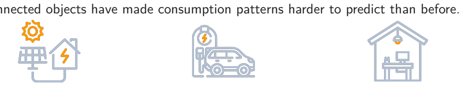
Abstract : This poster conceptually lays out recent advances in trustworthy machine learning (ML) that are of great interest for power systems applications like virtual power plants. All of them hint that ML, when done right, shouldn't be labelled as unreliable for critical applications where it could be beneficial. To be able to mitigate the recent uncertainty increase in power generation and demand, these tools are now crucial to plan the safe, efficient, and flexible operation of the grid while facilitating its decarbonization. The poster was presented at IVADO Digital Futures 2024 a popularization event on artificial intelligence and its applications.

(Trustworthy) AI for Québec's Virtual Power Plant

Julien Pallage, Antoine Lesage-Landry GERAD & MILA

1. Our current setting

Recent social changes have made it harder to predict the behaviour of electrical grids around the world. Our efforts in decarbonizing the grid, e.g., the addition of more intermittent renewable energy, the introduction of electrical vehicles on the road, and the integration of distributed energy resources (DERs), have generated a lot of uncertainty in the way we operate the grid. On the consumer side, the rising popularity of remote work and connected objects have made consumption patterns harder to predict than before.



(a) DERs (b) Electrical vehicles (c) Remote work

Figure 1: Some sources of uncertainty on the grid.

Virtual power plant (VPP) operators must resort to more sophisticated predictive algorithms, i.e., complex machine learning (ML), to plan the safe, efficient, and flexible operation of the grid while accounting for this added uncertainty [1].

2. Pattern recognition

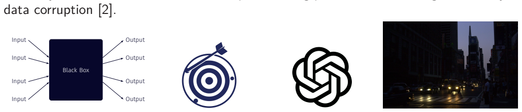
Machine Learning ≡ Algorithms that recognize patterns

(a) Neural Networks (Complex) (b) Linear Regression (Simple)

Figure 2: Two popular ML models of different complexity level

3. What's wrong with complex ML ?

Even though complex ML can recognize rich patterns in data to produce insightful predictions, they have been labelled as unreliable for critical sectors such as healthcare and energy. They tend to have a lack of interpretability, limited performance guarantees, a tendency to fail on never-seen data, complex training procedures, and a high sensitivity to data corruption [2].

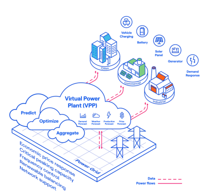


(a) Hard to interpret (b) No guarantees (c) Overconfidence (d) We hate blackouts!

Figure 3: Low deployment acceptability in critical sectors

4. Trustworthy ML for the virtual power plant

Fortunately, recent scientific advances have granted complex ML models more trustworthiness and, in the process, have made them more desirable for critical applications such as VPPs. VPPs are connected systems that reinforce grid flexibility through forecasting and control of DERs [3]. Since VPP operation relies on many predictions, e.g., weather, behaviour, and market, trustworthy ML models are the best of both worlds: they offer complex insights without compromising on reliability. Image: [4]



5. Respecting the laws of physics

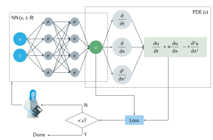
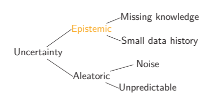


Figure 4: Architecture of a PINN [5]

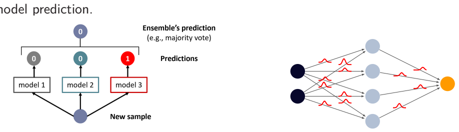
Would you have more faith in a model embedded with the physics of the application it is used in? Through numerous methods, researchers have made it possible to induce a physical intuition in ML models. As a result, these models usually require less training data, than their conventional counterparts, to perform as intended [5]. Physics-Informed Neural Networks (PINNs) are a popular example of physics-induced ML. In recent years, they have been used in countless applications, from fluid mechanics [6] to power systems modelling [7].

6. Learning the empirical uncertainty

Uncertainty as a whole is conceived as a combination of aleatoric and epistemic uncertainty. Simply put, aleatoric uncertainty accounts for what is purely random and can't be predicted, e.g., noise, while epistemic uncertainty reflects the uncertainty associated with missing knowledge [8].



By combining the prediction of multiple ML models, i.e., ensemble learning, and by building probabilistic predictive models, e.g., gaussian processes [9] and Bayesian Neural Networks (BNNs) [10], researchers have been able to capture the empirical uncertainty associated with model prediction.



(a) Ensembles [11] (b) Bayesian Neural Network

Figure 5: Two models that deal with epistemic uncertainty

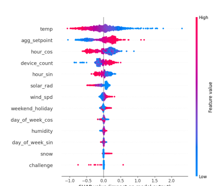
In other words, these models offer confidence intervals on their forecasts which can be useful for decision-making in critical sectors and applications, e.g., VPP operation.

7. Post-training verification

Since the training procedures of black box ML models is opaque and hard to understand, post-training verification frameworks have been developed to certify trained models' performance. It is comparable to a quality control (QC) procedure done on manufactured products before release. Some of these procedures can give theoretical worst-case violation guarantees for models used in constrained environments [12] or even theoretically bind the overall stability of a model [13].

8. Increased interpretability

Models whose decisions can't be well-interpreted or explained are hard to trust. As such, a lot of research has been done to add explainability to black-box models [14]. One of the most renowned approaches uses tools from game theory, i.e., Shapley values [15], to assign each input an importance value for any given prediction [16]. The figure presents a Shapley values analysis of a model predicting the hourly energy consumption of some Hilo users. Ask me how to interpret it!



9. Distributional robustness

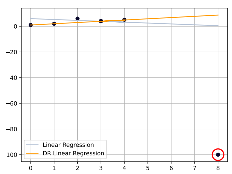


Figure 6: Regression with an outlier

ML models tend to be sensitive to outliers in the training data yet those outliers arise often when developing models for a VPP. Indeed, bad sensor readings, telemetry problems, numerical errors, as well as rare events, e.g., blackouts, snowstorms, and holidays, can corrupt the data used for training. Even though outlier filters are used to clean them out, it is impossible to guarantee their total disappearance from the dataset. Distributionally robust (DR) learning and optimization is a mathematical tool that can be leveraged to train ML models with empirical data that is assumed to be filled with outliers.

The intuition is that we aim to train our model to optimality on the most adverse dataset found in a "close range" from our currently available data [17, 18]. This generalized model, which is theoretically the best in the worst possible situation, has provable out-of-sample performance guarantees that mean a lot in VPP applications.

10. Our contribution

On this poster, we have conceptually laid out recent advances in trustworthy ML that we find very exciting. They all imply, in a way, that ML done right shouldn't be labelled as unreliable for critical applications. By using this as a starting point, our main goal is to develop new trustworthy ML models with theoretical and empirical guarantees for Québec's VPP.

To this date, we have developed a model that estimates the hourly peak-shaving capacity of Hilo users in localized demand response schemes. Also, we have been working on developing a trustworthy nonlinear baseline estimator. Stay tuned for what's next!

Special thanks to Salma, Bertrand, Ikhlas, Ahmed, Odile, and Steve from Hilo.

11. References

Articles and images:

[1] M. Khodayar, G. Liu, J. Wang, and M. E. Khodayar, "Deep learning in power systems research: A review," *CSEE Journal of Power and Energy Systems*, vol. 7, no. 2, pp. 209–220, 2021.
[2] K. Rasheed, A. Qayyum, M. Ghaly, A. Al-Fuqaha, A. Razi, and J. Qadir, "Explainable, trustworthy, and ethical machine learning for healthcare: A survey," *Computers in Biology and Medicine*, vol. 149, p. 106043, 2022.
[3] D. Pudjianto, C. Ramsay, and G. Strbac, "Virtual power plant and system integration of distributed energy resources," *IET Renewable Power Generation*, vol. 1, no. 1, p. 10, 2007.
[4] S. Klawiter, "What is a vpp?," Enel North America.
[5] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," *Nature Reviews Physics*, vol. 3, pp. 422–440, May 2021.
[6] Z. Mao, A. D. Jagtap, and G. E. Karniadakis, "Physics-informed neural networks for high-speed flows," *Computer Methods in Applied Mechanics and Engineering*, vol. 360, p. 112789, Mar. 2020.
[7] G. S. Misyris, A. Venzke, and S. Chatzivasileiadis, "Physics-informed neural networks for power systems," in *2020 IEEE Power Energy Society General Meeting (PESGM)*, pp. 1–5, 2020.
[8] A. D. Kiureghian and O. Ditlevsen, "Aleatory or epistemic? does it matter?," *Structural Safety*, vol. 31, no. 2, pp. 105–112, 2009. Risk Acceptance and Risk Communication.
[9] H. Liu, Y.-S. Ong, X. Shen, and J. Cai, "When gaussian process meets big data: A review of scalable gps," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 11, pp. 4405–4423, 2020.
[10] A. G. Wilson and P. Izmailov, "Bayesian deep learning and a probabilistic perspective of generalization," in *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS'20, (Red Hook, NY, USA), Curran Associates Inc., 2020.
[11] R. Yehoshua, "Introduction to ensemble methods," Towards AI, Aug 2023.
[12] A. Venzke, G. Qu, S. H. Low, and S. Chatzivasileiadis, "Learning optimal power flow: Worst-case guarantees for neural networks," *CoRR*, vol. abs/2006.11029, 2020.
[13] Y. Huang, H. Zhang, Y. Shi, J. Z. Kolter, and A. Anandkumar, "Training certifiably robust neural networks with efficient local lipschitz bounds," in *Advances in Neural Information Processing Systems* (M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan, eds.), vol. 34, pp. 22745–22757, Curran Associates, Inc., 2021.
[14] W. Samek, G. Montavon, S. Lapuschkin, C. J. Anders, and K.-R. Müller, "Explaining deep neural networks and beyond: A review of methods and applications," *Proceedings of the IEEE*, vol. 109, no. 3, pp. 247–278, 2021.
[15] L. S. Shapley, "A value for n-person games," in *Contributions to the Theory of Games II* (H. W. Kuhn and A. W. Tucker, eds.), pp. 307–317, Princeton: Princeton University Press, 1953.
[16] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems 30* (I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, eds.), pp. 4765–4774, Curran Associates, Inc., 2017.
[17] D. Kuhn, P. M. Esfahani, V. A. Nguyen, and S. Shafieezadeh-Abadeh, "Wasserstein distributionally robust optimization: Theory and applications in machine learning," OR Management Science in the Age of Analytics, 2019.
[18] P. M. Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations," Mathematical Programming, 2018.