Is learning for the unit commitment problem a low-hanging fruit?

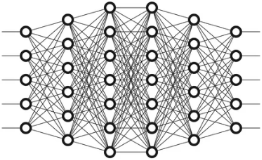
EURO 2022 (July 4-7 2022)

Salvador Pineda and Juan Miguel Morales

OASYS group, University of Málaga (Spain)

Machine learning in power systems

- The blast wave of machine learning has reached power systems
- Most papers propose black-box tools that run as follows:
  - Select a very hard problem to solve (usually NP-hard)
  - Set hyperparameters $\mu, \epsilon_0, \alpha, \beta, \lambda, \ldots$ (without explaining how)
  - Train a deep neural network (using available software)



  - Discuss the computational savings of the proposed learning-based method with that of solving the original problem

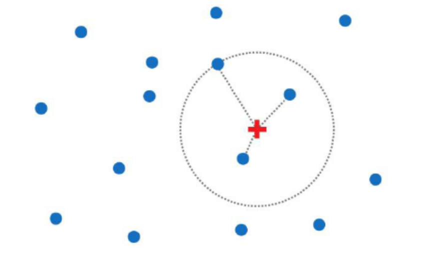
Questions we want to address

- Are power system problems so complex that any fortuitous learning-based method may involve significant computational savings?
- Do simpler and interpretable learning-based methods perform similarly to complex black-box methods?
- Should black-box methods be benchmarked against simpler methods in power systems applications?

Unit commitment problem

Unit Commitment Problem	
Horizon	24 hours
Obj	Min production cost
Var	On/off status (binary) Generating dispatches
Con	Generation = Demand Unit technical limits Line technical limits

Nearest Neighbors



AIM
Leverage past unit commitment solutions to solve new instances of the problem

Unit commitment formulation

$$\min_{\mathbf{x}\in\mathbb{R}^n,\, \mathbf{y}\in\{0,1\}^m} \quad f(\mathbf{x},\mathbf{y}) \tag{1a}$$

$$g_i(\mathbf{x},\mathbf{y}) \leqslant 0, \quad \forall i \tag{1b}$$

$$h_j(\mathbf{x},\mathbf{d}) \leqslant 0, \quad \forall j \tag{1c}$$

- Varying input parameters $\mathbf{d}$: demand, renewable power generation
- Continuous variables $\mathbf{x}$: power dispatches, power flows through lines
- Binary variables $\mathbf{y}$: on/off status of the generating units
- Objective function (1a) minimizes the total generation costs
- Equation (1b): technical constraints of generating units
- Equation (1c): technical constraints of network
- Even if all functions are linear, problem (1a)-(1c) is **NP-hard**

Historical data

We have access to a set of historical data including:

- Input parameters $\{\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_N\}$
- Optimal continous decisions $\{\mathbf{x}_1^*, \mathbf{x}_2^*, \ldots, \mathbf{x}_N^*\}$
- Optimal binary decisions $\{\mathbf{y}_1^*, \mathbf{y}_2^*, \ldots, \mathbf{y}_N^*\}$

KNN + LP approach

For a new input vector $\tilde{\mathbf{d}}$ do:

- Among $\{\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_N\}$, find the K nearest neigborst to $\tilde{\mathbf{d}}$.
- For each neighbor k do
  - Fix binary variables to $\mathbf{y}_k^*$
  - Solve the linear program

$$\min_{\mathbf{x}\in\mathbb{R}^n} \quad f(\mathbf{x}, \mathbf{y}_k^*) \tag{2a}$$

$$g_i(\mathbf{x}, \mathbf{y}_k^*) \leqslant 0, \quad \forall i \tag{2b}$$

$$h_j(\mathbf{x}, \tilde{\mathbf{d}}) \leqslant 0, \quad \forall j \tag{2c}$$

  - Denote the optimal solution and optimal value as $\mathbf{x}_k^L$ and z_k^L
- Choose the neighbor with minimum cost $\tilde{k} \in \arg\min_k \{z_k^L\}$
- Provide the optimal solution $\mathbf{x}_{\tilde{k}}^L$ and $\mathbf{y}_{\tilde{k}}^*$

KNN + LP results

- 500 instances, leave-one-out
- $K = 50$
- MIP gap = 0.01%

System	Av. error (%)	Max error (%)	$< 0.01\%$	# Infes	Speedup
1888rte	0.0174	0.2394	230	1	116.5x
1951rte	0.0382	0.3759	47	8	150.4x
2848rte	0.0186	0.1332	179	2	132.6x
3012wp	0.0485	0.4864	37	5	188.8x
3375wp	0.1256	0.8073	9	13	215.9x
6468rte	-0.0001	0.0175	498	0	41.2x
6470rte	-0.0016	0.0187	496	0	171.9x
6495rte	-0.0001	0.0481	496	0	41.0x
6515rte	-0.0009	0.0133	497	0	101.7x

KNN + LP results

System	Av. error (%)	Max error (%)	$< 0.01\%$	# Infes	Speedup
1888rte	0.0174	0.2394	230	1	116.5x
1951rte	0.0382	0.3759	47	8	150.4x
2848rte	0.0186	0.1332	179	2	132.6x
3012wp	0.0485	0.4864	37	5	188.8x
3375wp	0.1256	0.8073	9	13	215.9x
6468rte	-0.0001	0.0175	498	0	41.2x
6470rte	-0.0016	0.0187	496	0	171.9x
6495rte	-0.0001	0.0481	496	0	41.0x
6515rte	-0.0009	0.0133	497	0	101.7x

- No need for complicated learning techniques for these systems, as naive strategies involve time reductions with negligible errors.
- For these systems, the naive learning strategy involves errors slightly higher than the set MIP gap, but with substantial time reductions.
- For these systems, the naive approach involves higher errors and some infeasible cases. Thus, other learning approaches may be required.

KNN + LP results

System	Av. error (%)	Max error (%)	$< 0.01\%$	# Infes	Speedup
1888rte	0.0174	0.2394	230	1	116.5x
1951rte	0.0382	0.3759	47	8	150.4x
2848rte	0.0186	0.1332	179	2	132.6x
3012wp	0.0485	0.4864	37	5	188.8x
3375wp	0.1256	0.8073	9	13	215.9x
6468rte	-0.0001	0.0175	498	0	41.2x
6470rte	-0.0016	0.0187	496	0	171.9x
6495rte	-0.0001	0.0481	496	0	41.0x
6515rte	-0.0009	0.0133	497	0	101.7x

- No need for complicated learning techniques for these systems, as naive strategies involve time reductions with negligible errors.
- For these systems, the naive learning strategy involves errors slightly higher than the set MIP gap, but with substantial time reductions.
- For these systems, the naive approach involves higher errors and some infeasible cases. Thus, other learning approaches may be required.

KNN + LP results

System	Av. error (%)	Max error (%)	$< 0.01\%$	# Infes	Speedup
1888rte	0.0174	0.2394	230	1	116.5x
1951rte	0.0382	0.3759	47	8	150.4x
2848rte	0.0186	0.1332	179	2	132.6x
3012wp	0.0485	0.4864	37	5	188.8x
3375wp	0.1256	0.8073	9	13	215.9x
6468rte	-0.0001	0.0175	498	0	41.2x
6470rte	-0.0016	0.0187	496	0	171.9x
6495rte	-0.0001	0.0481	496	0	41.0x
6515rte	-0.0009	0.0133	497	0	101.7x

- No need for complicated learning techniques for these systems, as naive strategies involve time reductions with negligible errors.
- For these systems, the naive learning strategy involves errors slightly higher than the set MIP gap, but with substantial time reductions.
- For these systems, the naive approach involves higher errors and some infeasible cases. Thus, other learning approaches may be required.

KNN + MIP approach

For a new input vector $\tilde{\mathbf{d}}$ do:

- Among $\{\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_N\}$, find the K nearest neigborst to $\tilde{\mathbf{d}}$.
- Set upper bound of $\mathbf{y}$ to $\lceil \frac{1}{K} \sum_k \mathbf{y}_k^* \rceil$.
- Set lower bound of $\mathbf{y}$ to $\lfloor \frac{1}{K} \sum_k \mathbf{y}_k^* \rfloor$.
- Solve the mixed-integer linear program

$$\min_{\mathbf{x} \in \mathbb{R}^n, \, \mathbf{y} \in \{0,1\}^m} \quad f(\mathbf{x}, \mathbf{y}) \tag{3a}$$

$$g_i(\mathbf{x}, \mathbf{y}) \leqslant 0, \quad \forall i \tag{3b}$$

$$h_j(\mathbf{x}, \tilde{\mathbf{d}}) \leqslant 0, \quad \forall j \tag{3c}$$

$$\lfloor \frac{1}{K} \sum_k \mathbf{y}_k^* \rfloor \leqslant \mathbf{y} \leqslant \lceil \frac{1}{K} \sum_k \mathbf{y}_k^* \rceil \tag{3d}$$

KNN + MIP results

System	Av. error (%)	Max error (%)	$< 0.01\%$	# Infes	Speedup
1888rte	-0.0004	0.0141	499	0	30.7x
1951rte	0.0002	0.0227	499	0	7.7x
2848rte	0.0004	0.0228	497	0	20.0x
3012wp	-0.0001	0.0199	497	0	15.4x
3375wp	0.0001	0.0198	494	0	14.6x
6468rte	-0.0002	0.0217	499	0	18.9x
6470rte	0.0007	0.0174	499	0	15.9x
6495rte	-0.0001	0.0086	500	0	8.5x
6515rte	0.0003	0.0151	498	0	26.5x

- Most cases below GAP
- No infeasible cases
- Speedup factor between 7.7x and 30.7x

Questions we want to address

- Are power system problems so complex that any fortuitous learning-based method may involve significant computational savings?

  Yes, specially in combinatorial problems like UC

- Do simpler and interpretable learning-based methods perform similarly to complex black-box methods?

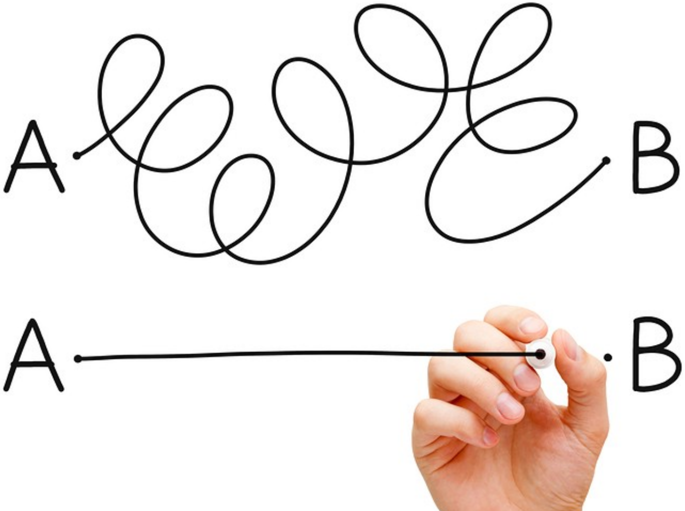
  Yes, depending on problem structure and data

- Should black-box methods be benchmarked against simpler methods in power system applications?

  Yes. Otherwise, irrelevant publications will continue

Takeaway message

In many cases the simplicity of less is more



Thanks for the attention!! Questions??

Electric Power Systems Research 207 (2022) 107851

Contents lists available at ScienceDirect

Electric Power Systems Research

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

Is learning for the unit commitment problem a low-hanging fruit?

S. Pineda *, J.M. Morales

University of Malaga, Spain

Optimization and Analytics for Sustainable energY Systems

More info: oasys.uma.es

Email: spineda@uma.es