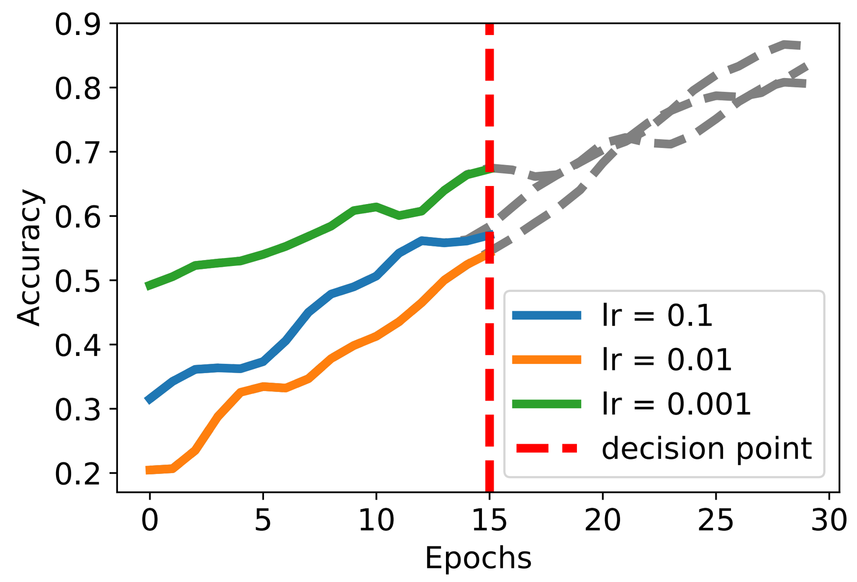
A Hybrid Quantum-Classical Workflow for Hyperparameter Optimization of Neural Networks

Marcel Aach[1,2], Eric Wulff[3], Edoardo Pasetto[1], Amer Delilbasic[1,2], Rakesh Sarma[1], Eray Inanc[1], Maria Girone[3], Morris Riedel[2], Andreas Lintermann[1]

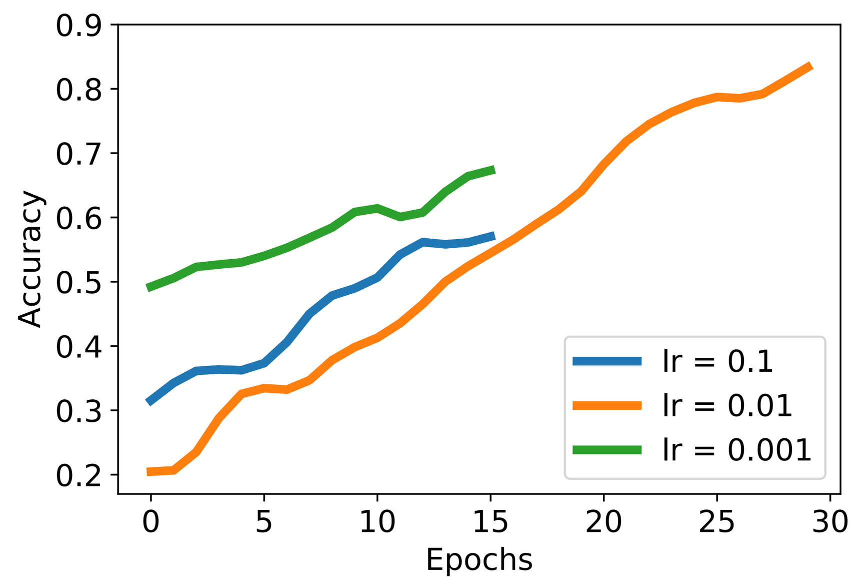
Jülich Supercomputing Centre[1], University of Iceland[2], CERN[3]

1. PROBLEM DEFINITION

- HPO requires significant computational resources
- Run trials on classical GPUs
- Early stopping of bad trials with Quantum Annealer (QA) by Quantum Support Vector Regression (Q-SVR) learning curve extrapolation



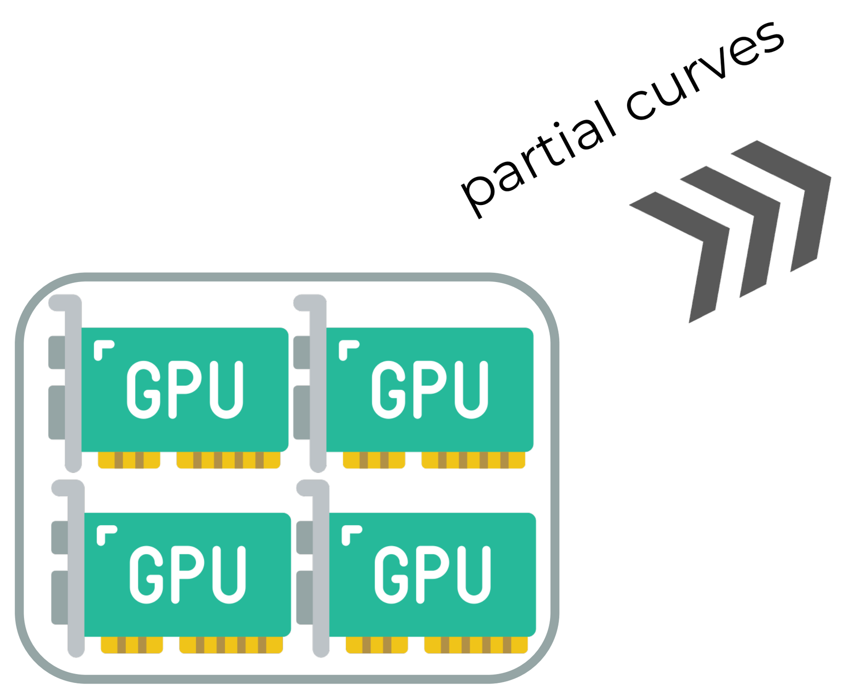
Good vs. bad trial?

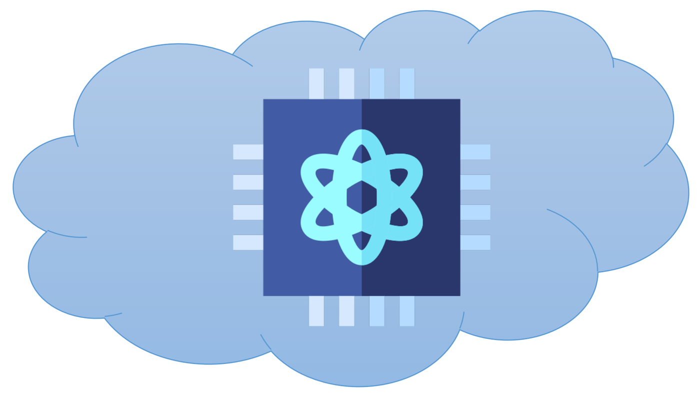


Only continue good trials!

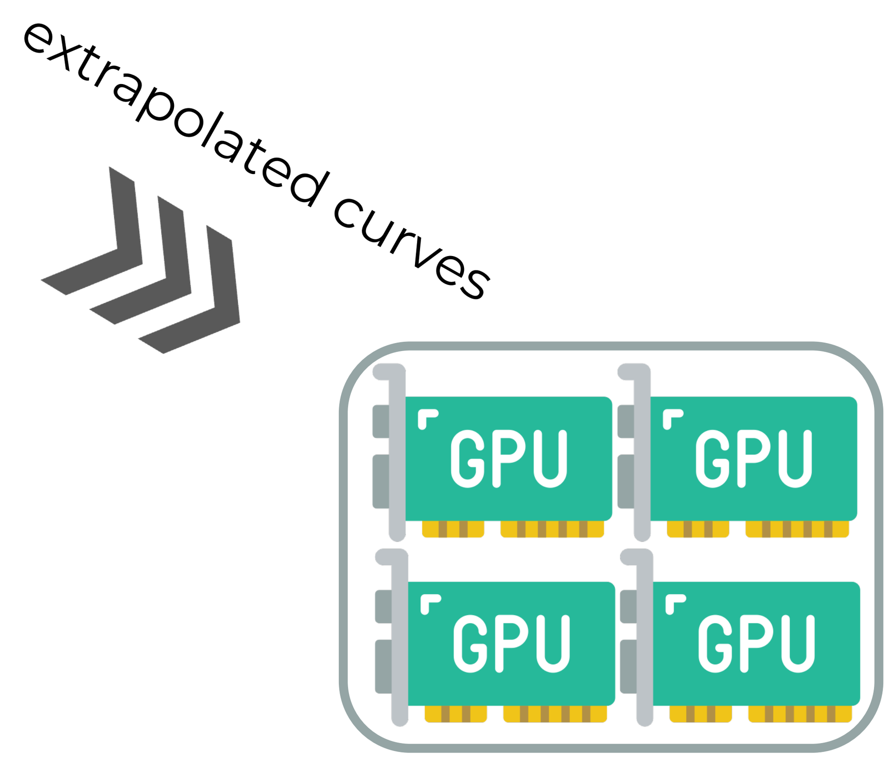
3. WORKFLOW

Three step workflow





Extrapolate learning curves with Q-SVR on the *Quantum Annealer* via D-Wave cloud



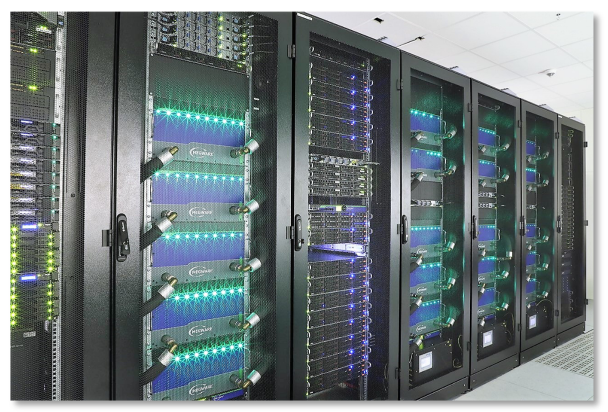
① Partial neural network training on *classical supercomputer*

② Extrapolate learning curves with Q-SVR on the *Quantum Annealer* via D-Wave cloud

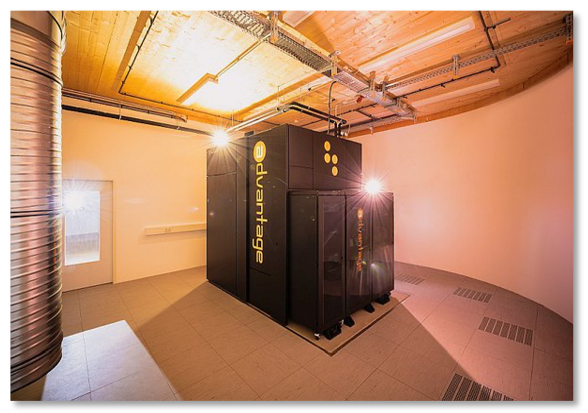
③ Full training of best neural networks on *classical supercomputer*

Systems

DEEP-EST-ESB



JUPSI

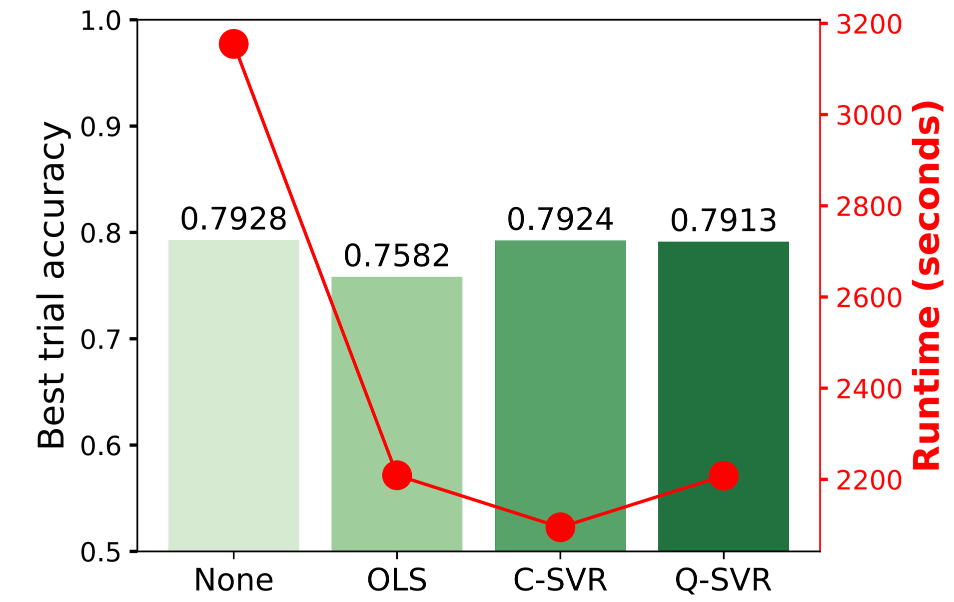


Hyperparameter Search Space

Convolutional neural network trained on cifar-10

Parameter	Range
Layers	[2, 3, 4]
Filters	[16, 32, 48, 64]
Batch size	[64, 128, 256, 512]
Learning rate	Log[1e-4, 1]
Momentum	Log[1e-4, 0.9]

4. RESULTS



- 100 trials with different early stopping methods
- No early stopping: longest runtime but highest accuracy
- Q-SVR: 30% shorter runtime while comparable accuracy

Why Quantum Support Vector Regression?

- C-SVR complexity $O(n^2)$ - $O(n^3)$
- Q-SVR complexity does not increase with problem size
- Quantum speedup possible in the future

2. THEORETICAL BACKGROUND

Performance Prediction:

- Regression on learning curve of a neural network can accurately predict final performance [Baker et al. 2018]

Quantum Support Vector Regression:

- QA solves quadratic unconstrained binary optimization (QUBO) problems: $E = \sum_{i \leq j} a_i Q_{ij} a_j, \quad a_{i,j} \in \{0,1\}$
- Reforumulate Classical Support Vector Regression (C-SVR) into QUBO form [Pasetto et al. 2022]:

Kernel function

$$L(\boldsymbol{\alpha}, \hat{\boldsymbol{\alpha}}) = \frac{1}{2}\sum_{n=0}^{N-1}\sum_{m=0}^{N-1}(\alpha_n - \hat{\alpha}_n)(\alpha_m - \hat{\alpha}_m)k(\boldsymbol{x}_n, \boldsymbol{x}_m) - \epsilon\sum_{n=0}^{N-1}(\alpha_n + \hat{\alpha}_n) + \sum_{n=0}^{N-1}(\alpha_n - \hat{\alpha}_n)y_n$$

Error margin

Number of qubits

$$\alpha_n = \sum_{k=0}^{K-1} B^{k-P} a_{Kn+k}$$

Encoding basis

$$\hat{\alpha}_n = \sum_{k=0}^{K-1} B^{k-P} a_{K(N+n)+k}$$

5. CONCLUSION

- Neural network training on classical machine
- Performance prediction on quantum machine
- Save 30% compute resources

REFERENCES

- Baker, B., Gupta, O., Raskar, R., & Naik, N.. (2017). Accelerating Neural Architecture Search using Performance Prediction.
- E. Pasetto, M. Riedel, F. Melgani, K. Michielsen and G. Cavallaro, "Quantum SVR for Chlorophyll Concentration Estimation in Water With Remote Sensing," in IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-5, 2022, Art no. 1505705, doi: 10.1109/LGRS.2022.3200325.
- Suarez, E.; Kreuzer, A.; Eicker, N.; Lippert, T. The DEEP-EST project. In *Porting Applications to a Modular Supercomputer-Experiences from the DEEP-EST Project*; Schriften des Forschungszentrums Jülich IAS Series; Forschungszentrum Jülich GmbH Zentralbibliothek, Verlag: Jülich, Germany, 2021; Volume 48, pp. 9–25.
- Images taken from: https://www.fz-juelich.de/en/ias/jsc/systems/prototype-systems/deep_system and https://www.fz-juelich.de/en/ias/jsc/systems/quantum-computing/juniq-facility/juniq/d-wave-advantagetm-system-jupsi

- This work was performed in the Center of Excellence (CoE) Research on AI- and Simulation-Based Engineering at Exascale (RAISE) receiving funding from EU's Horizon 2020 Research and Innovation Framework Programme H2020-INFRAEDI-2019-1 under grant agreement no. 951733.
- The authors gratefully acknowledge the Jülich Supercomputing Centre for funding this project by providing computing time on the D-Wave Advantage™ System JUPSI through the Jülich Unified Infrastructure for Quantum computing (JUNIQ).