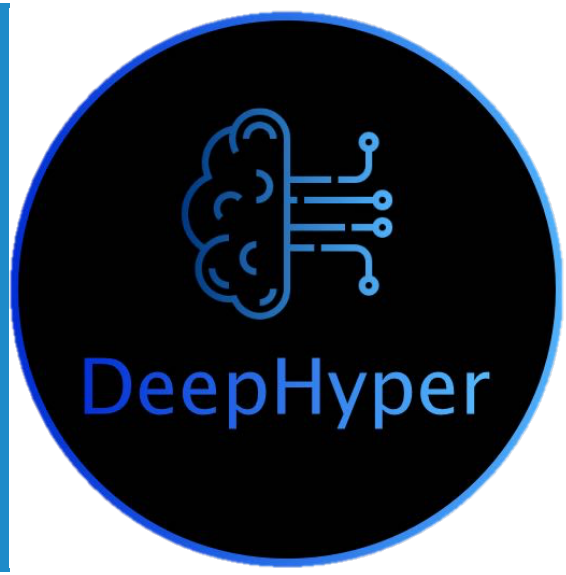
MARCH 9, 2022

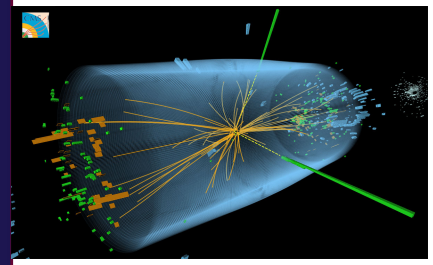
Argonne NATIONAL LABORATORY 75 1946–2021

DeepHyper: Automated Machine Learning at Scale

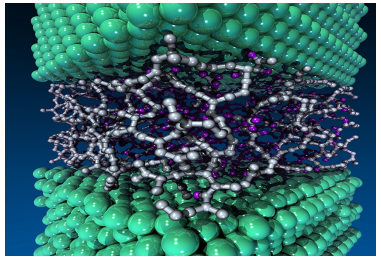


Prasanna Balaprakash
Computer Scientist
Mathematics & Computer Science Division
Leadership Computing Facility
Argonne National Laboratory
pbalapra@anl.gov
https://www.mcs.anl.gov/~pbalapra/

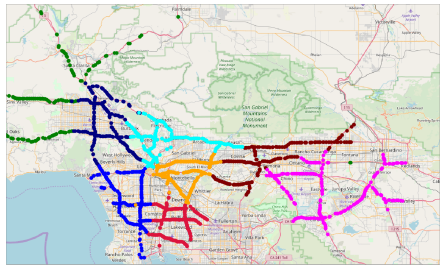
AI/ML for DOE applications



Subatomic Particles



Molecular Dynamics

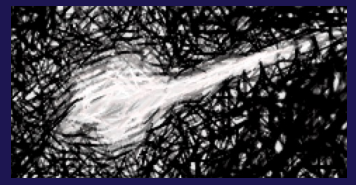


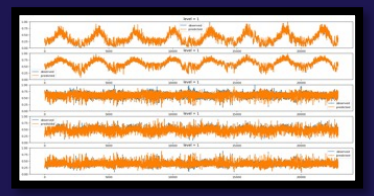
Transportation

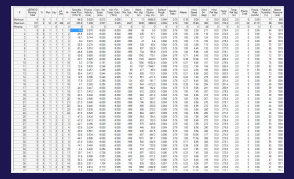
Weather

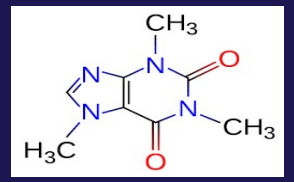
Cosmology

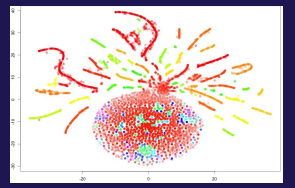
More than 100+ AI/ML applications at ANL in 2018 (large tail)

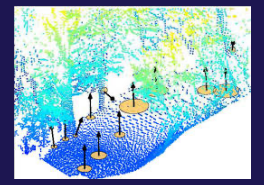






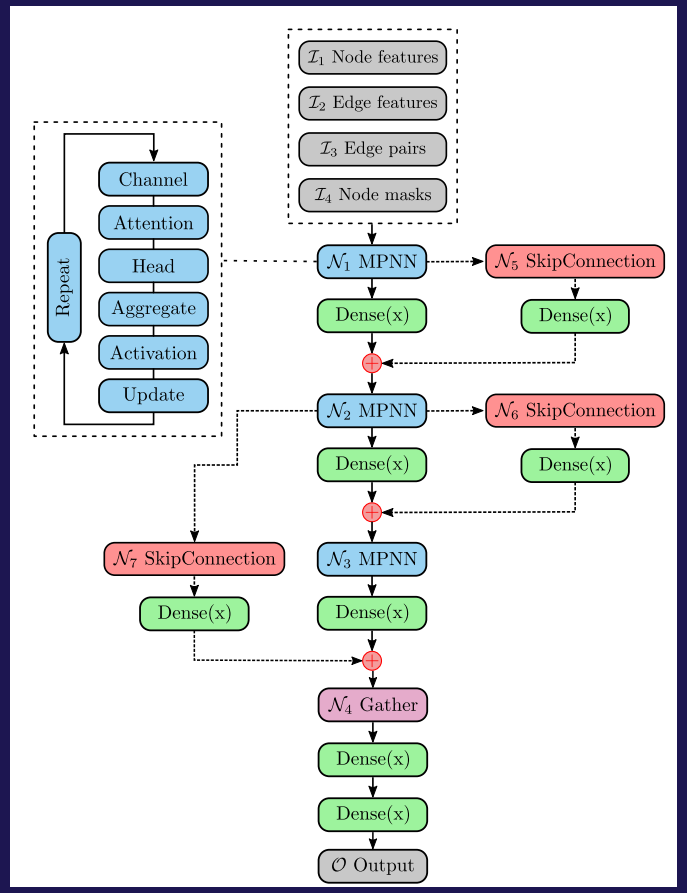






Diverse data types

Degrees of Freedom in Neural Networks Design

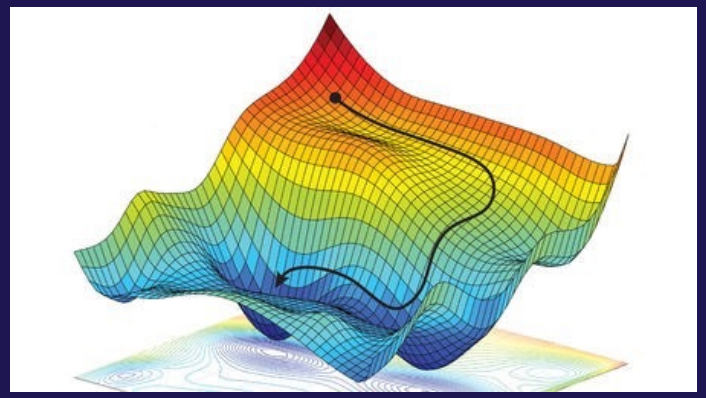


~10^{10} possible designs

Manual design: trial and error and time consuming (aka *graduate student descent*)

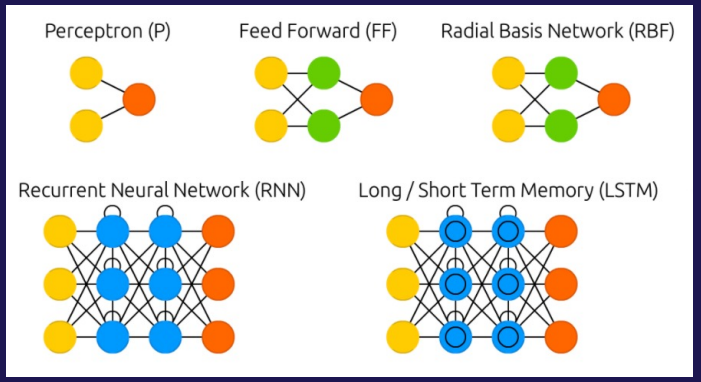
Degrees of Freedom in Neural Networks Design

Algorithm Hyperparameters



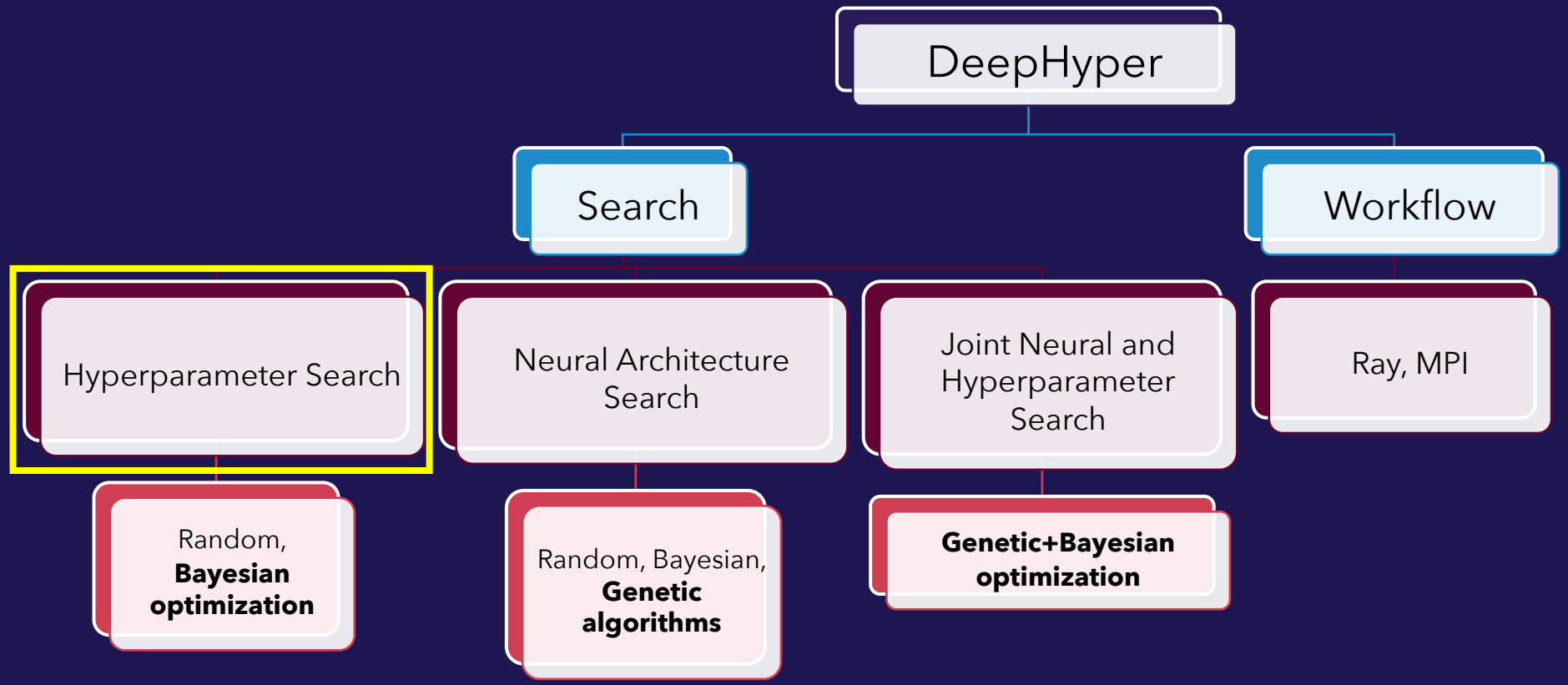
Optimizer: SGD, RMSprop, Adam…
Learning rate
Minibatch size
Learning rate scheduler
Adaptative batch size
…

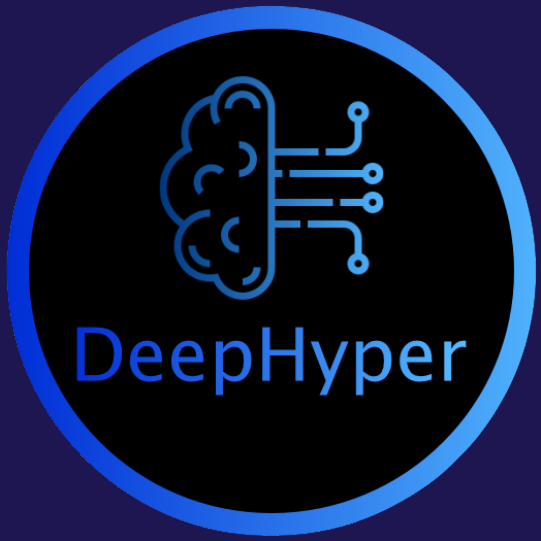
Architecture Variables



Number of layers
Type of the layer: Fully Connected, Convolution, Recursive…
Activation function
Dropout rate
Skip connection
…

DeepHyper: Scalable AutoML



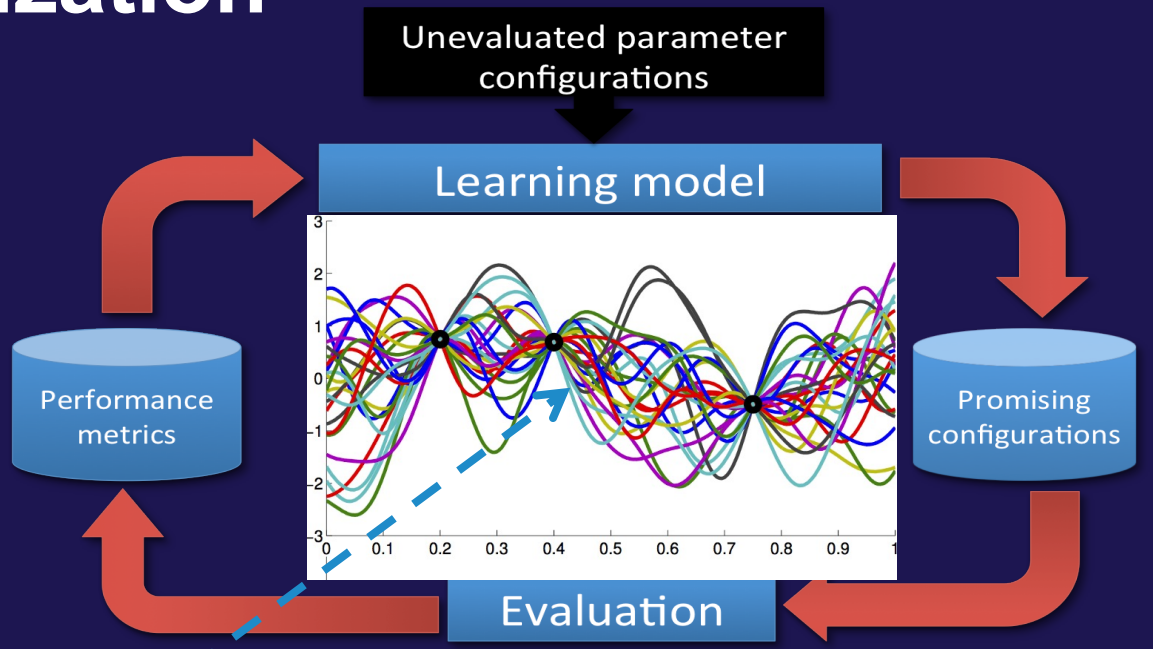


DeepHyper documentation: http://deephyper.readthedocs.io

Asynchronous Bayesian Optimization

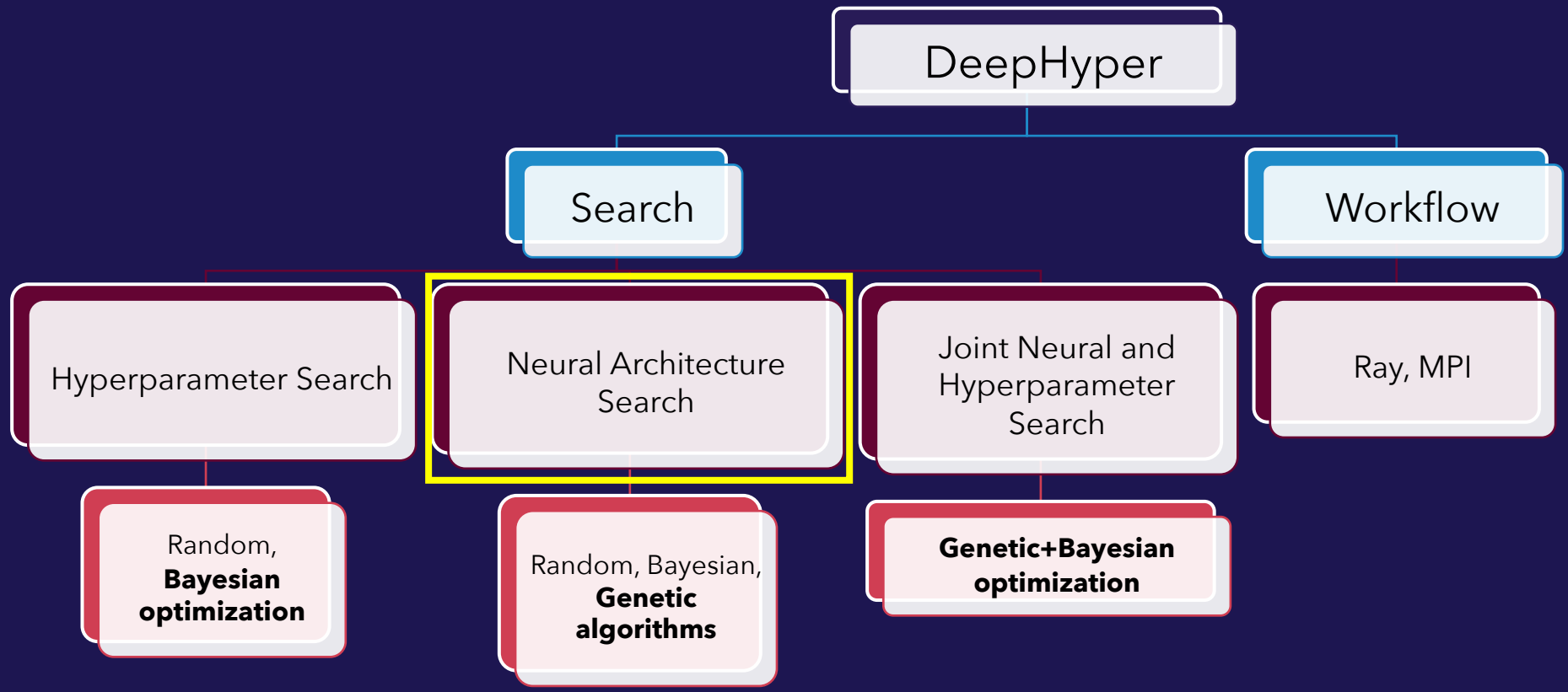
Algorithm: Asynchronous BO

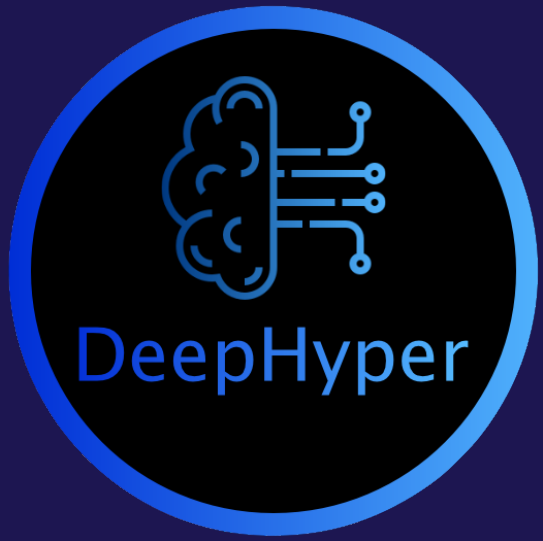
```
   /* Initialization                                          */
1  optimizer ← optimizer()
2  for i ← 1 to W do
3  |  configs.h_m ← random_point(H_m)
4  |  submit_evaluation(configs) // Nonblocking
5  end
   /* Main loop                                               */
6  while not done do
   |  // Query results
7  |  results ← get_finished_evaluations ()
8  |  if |results| > 0 then
9  |  |  optimizer.tell(results.h_m, results.valid_accuracy)
10 |  |  next ← optimizer.ask(|results|)
11 |  |  submit_evaluation(next) // Nonblocking
12 |  end
13 end
```



Surrogate Model Fitted to Sampled Performance (iterative refinement improves the learning model)

DeepHyper: Scalable AutoML





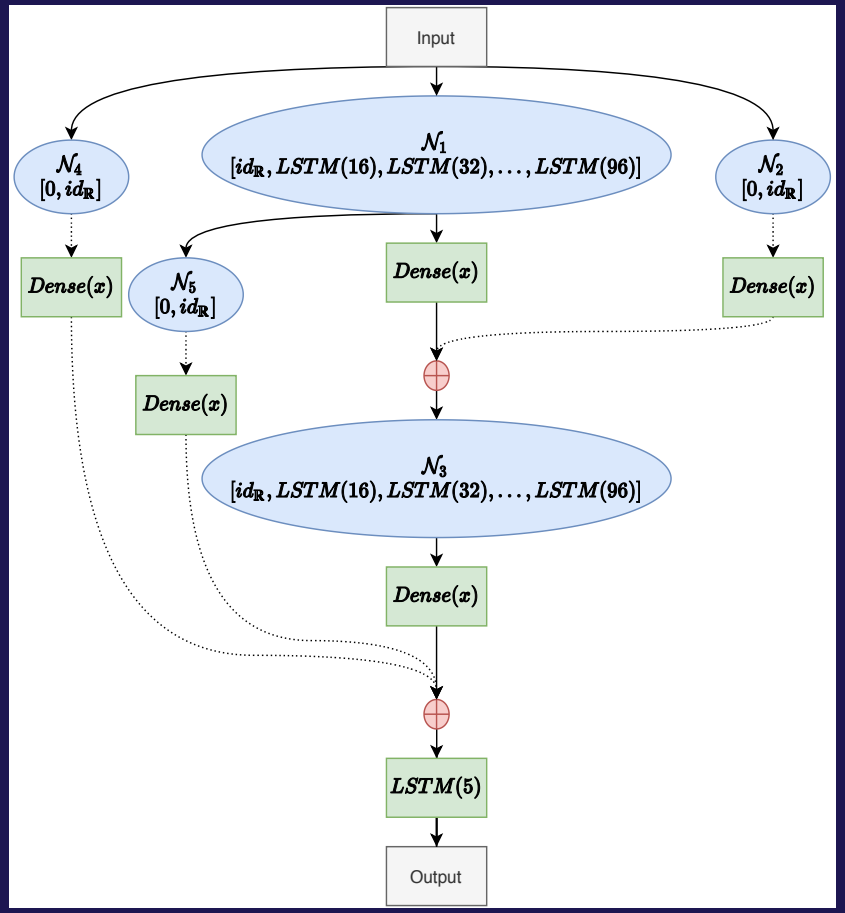
DeepHyper documentation: http://deephyper.readthedocs.io

Configuring Neural Architecture Search

How do we define a space of neural networks?

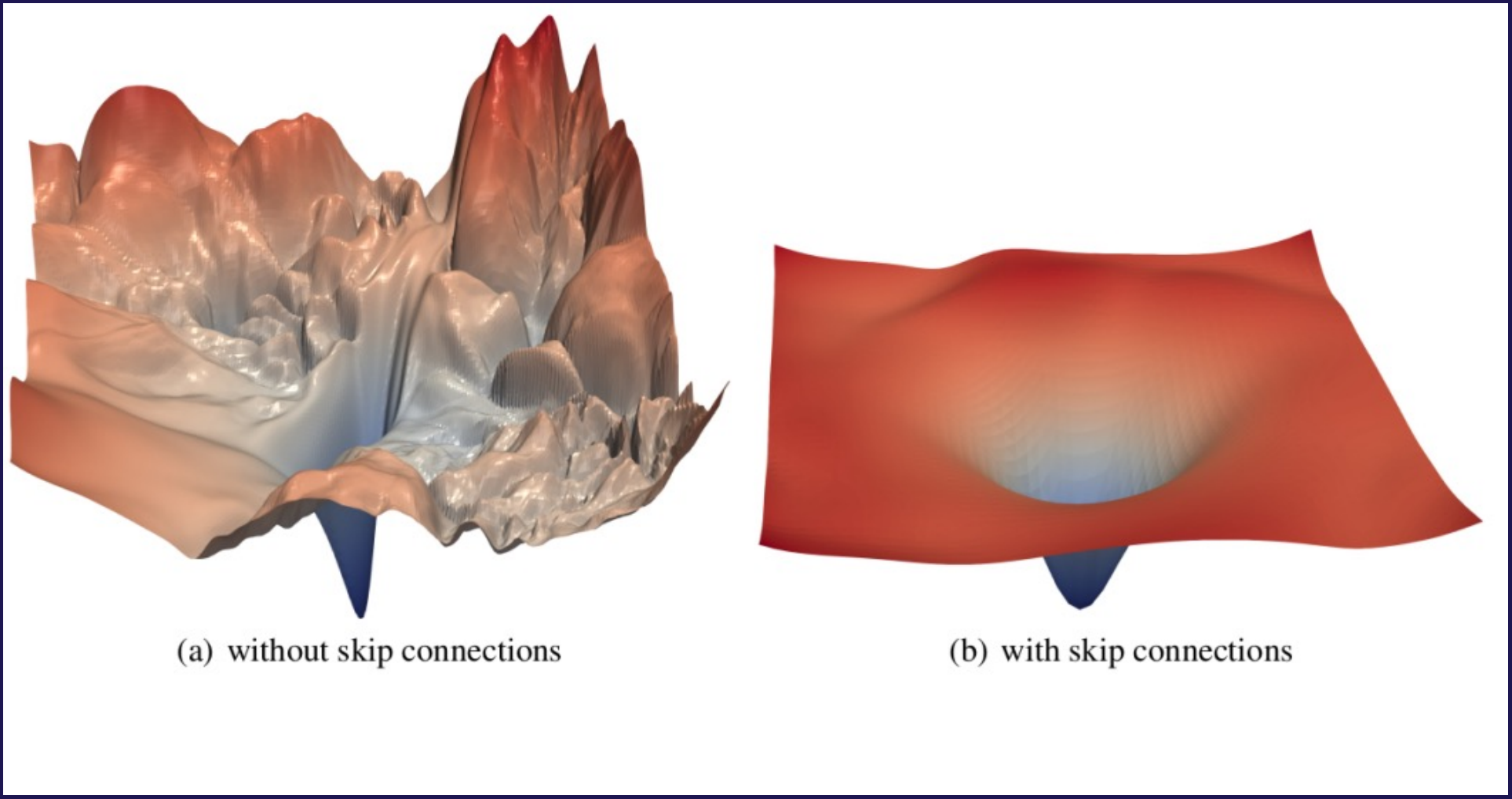
A neural network search space can be represented as a directed acyclic graph with nodes and edges.

Nodes represent possible operations, for example:
1. Add an identity layer
2. Add a layer with 40 neurons
3. Add a layer with 60 neurons
4. Add a dropout operation
5. Add a skip connection to another node



- Nodes can be constant – (i.e., predefined and immutable during the search)
- Nodes can be variable – (i.e., the search can tweak these to get better performance)
- Each variable node has an upper bound on the number of operations (which may be expressed as a categorical variable). Edges define the flow of the tensor in the graph

Skip Connections



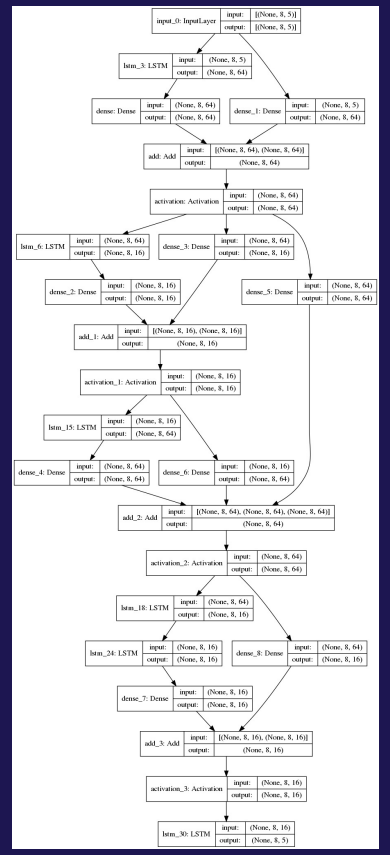
Loss surface: with and without skip connections
Physics-informed neural networks

Li, Hao, Zheng Xu, Gavin Taylor, Christoph Studer, and Tom Goldstein. "Visualizing the loss landscape of neural nets." Advances in neural information processing systems 31 (2018).
A. Krishnapriyan, A. Gholami, S. Zhe, R. Kirby, and M. W. Mahoney. "Characterizing possible failure modes in physics-informed neural networks." Advances in Neural Information Processing Systems 34 (2021).

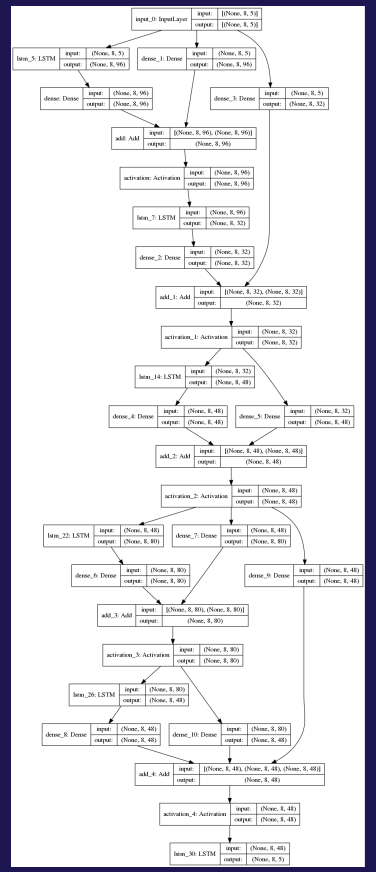
DeepHyper NAS-API

```
search_space = create_search_space(num_layers=5)
ops = [random() for _ in range(search_space.num_nodes)]
search_space.set_ops(ops)
model = search_space.create_model()
model.summary()
plot_model(model, to_file='sampled_neural_network.png', show_shapes=True)
print("The sampled_neural_network.png file has been generated.")
```

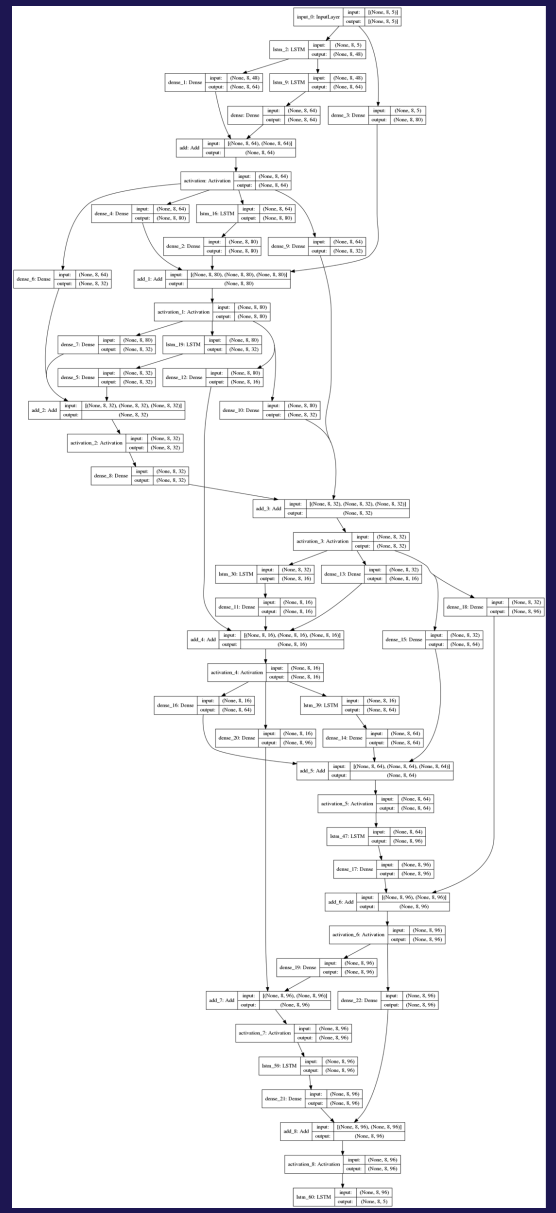
Fun to generate random architectures!



68,152 parameters



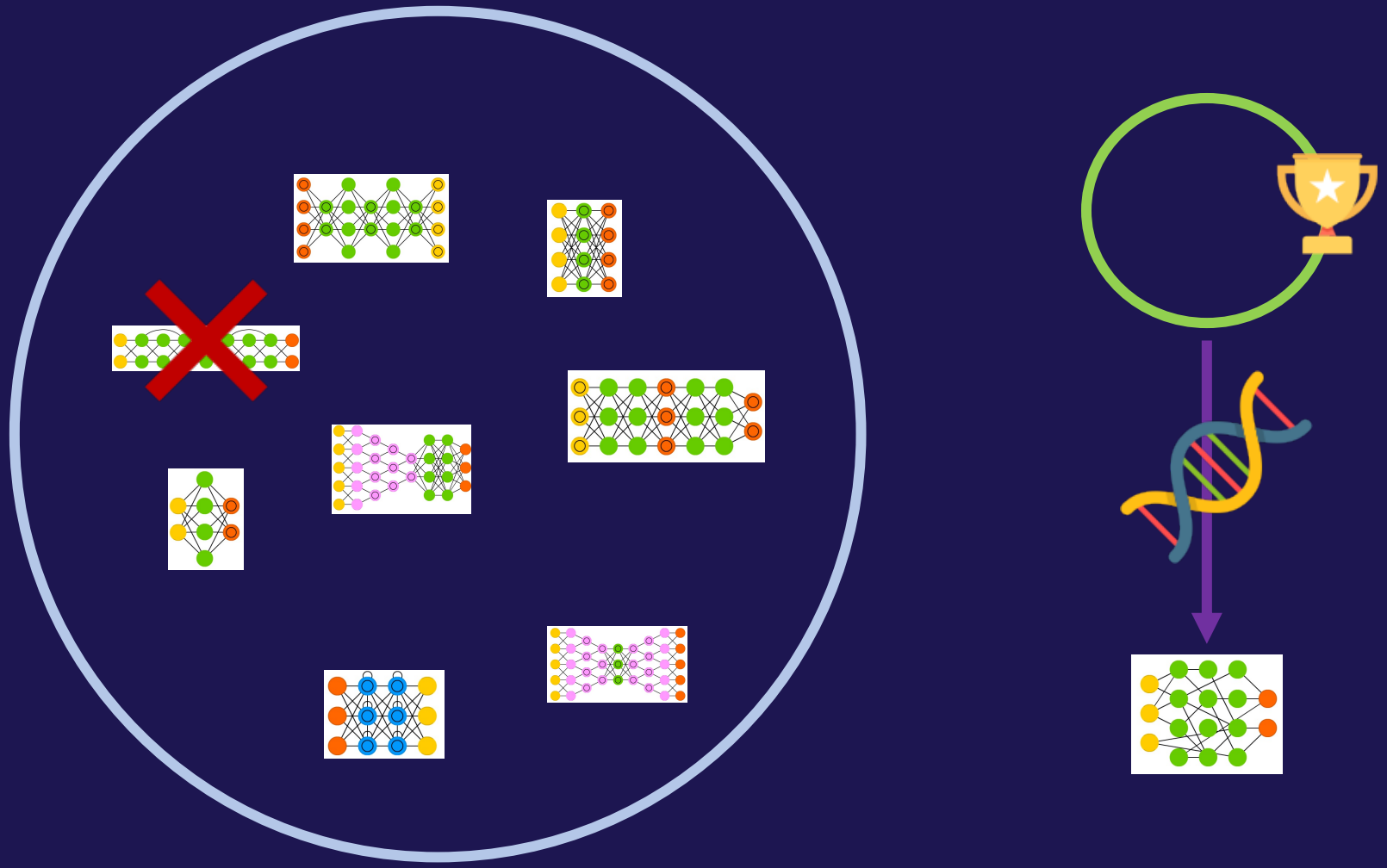
172,424 parameters



344,424 Parameters (more skips/layers)

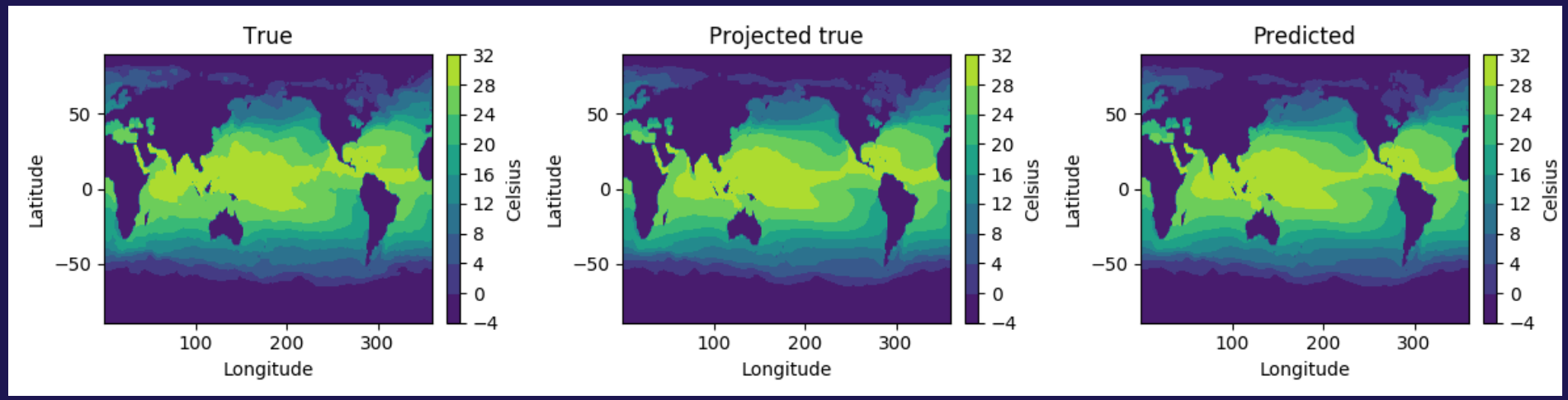
Exploring Search Space

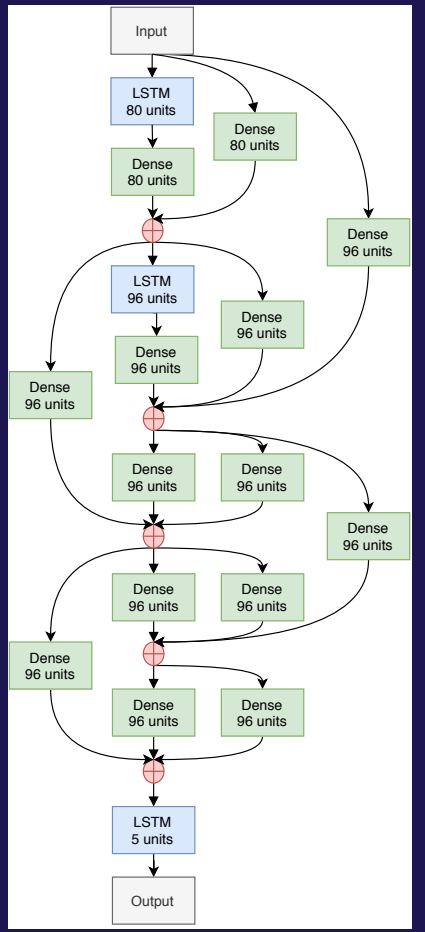
Regularized ageing evolution to explore the search space of possible architectures



Real, Esteban, et al. "Regularized evolution for image classifier architecture search." Proceedings of the AAAI conference on artificial intelligence. Vol. 33. 2019.

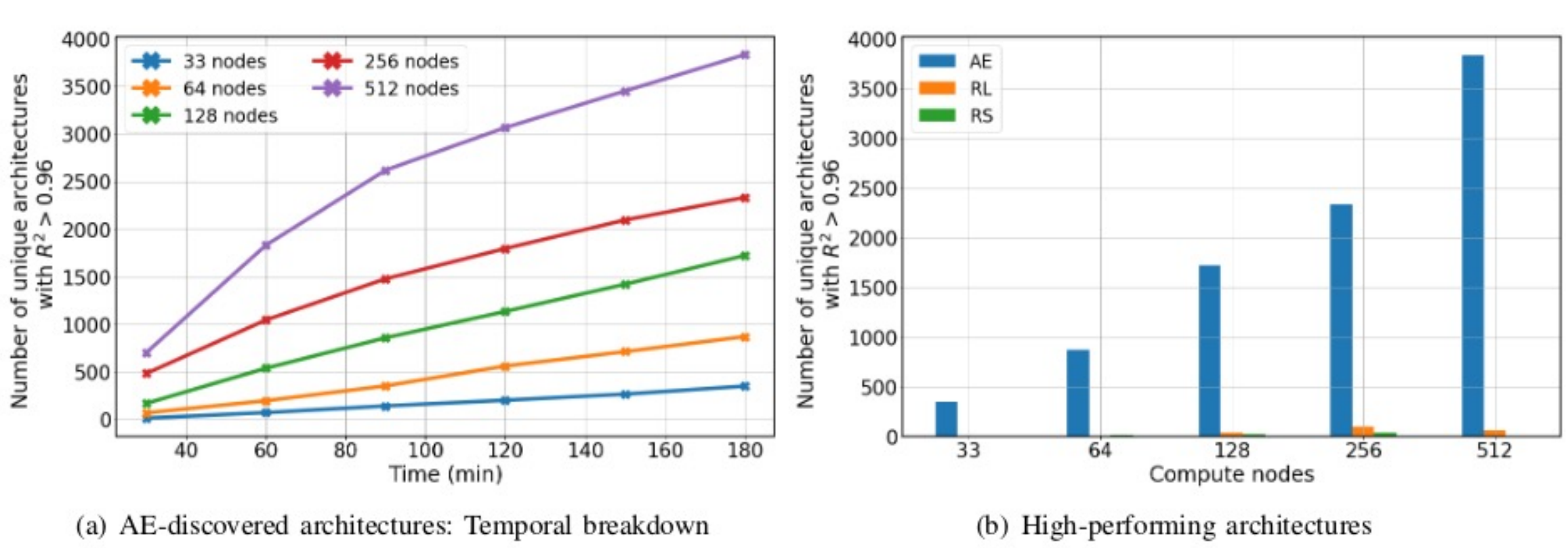
Searching for a Surrogate LSTM: Sea Surface Temperature Forecasting





R. Maulik, R. Egele, B. Lusch, and P. Balaprakash. Recurrent Neural Network Architecture Search for Geophysical Emulation. In SC '20: IEEE/ACM International Conference on High Performance Computing, Networking, Storage and Analysis, 2020.

Searching for a Surrogate LSTM: Sea Surface Temperature Forecasting



(a) AE-discovered architectures: Temporal breakdown

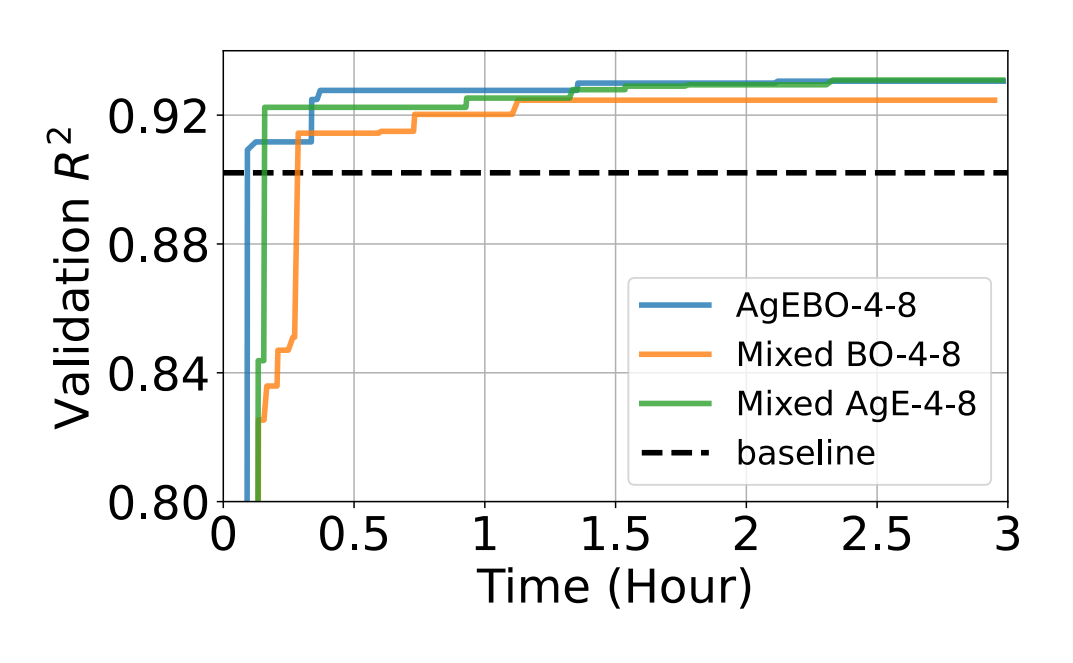
(b) High-performing architectures

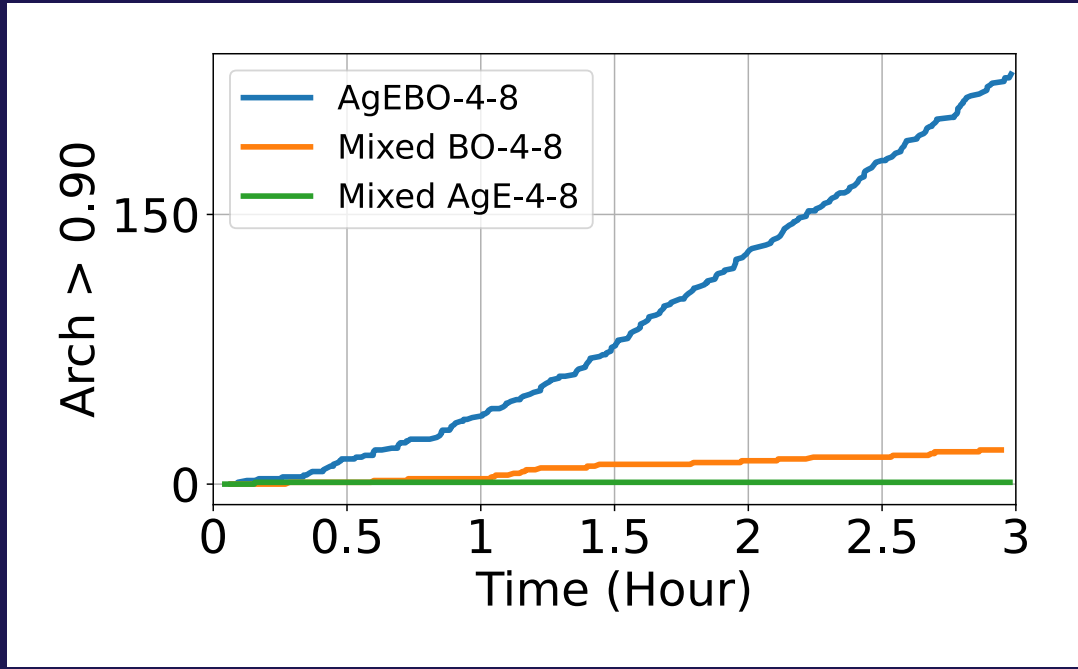
ALCF/Theta

R. Maulik, R. Egele, B. Lusch, and P. Balaprakash. Recurrent Neural Network Architecture Search for Geophysical Emulation. In SC '20: IEEE/ACM International Conference on High Performance Computing, Networking, Storage and Analysis, 2020.

Cancer Drug Response:
predicting the activity of a drug treatment against a cancer cell

Method–GPU(s)/Eval–Nodes





ALCF/ThetaGPU

R. Egele, P. Balaprakash, I. Guyon, V. Vishwanath, F. Xia, R. Stevens, Z. Liu. AgEBO-tabular: joint neural architecture and hyperparameter search with autotuned data-parallel training for tabular data. In SC '21: IEEE/ACM International Conference on High Performance Computing, Networking, Storage and Analysis, 2021.

Active Research Topics

- AutoML at (exa-)scale
- Deep ensembles and uncertainty quantification
- Multiobjective optimization (accuracy, inference time, #params)
- Designing insect-brain inspired learning algorithms for neuromorphic computing
- Software/hardware co-design, compiler optimization, data services management
- Continual/transfer learning across similar tasks

The DeepHyper Community


Romain Egele


Misha Salim


Stefan Wild


Venkatram Vishwanath


Romit Maulik


Kyle Gerard Felker

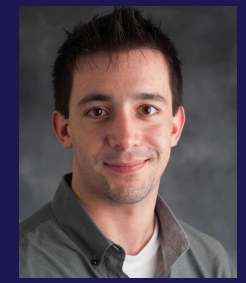

Taylor Childers


Tom Uram


Matthieu Dorier


Sandeep Madireddy


Bethany Lusch


Sam Foreman


Shengli Jiang


Mansi Sakarvadia


Elise Jennings


Tanwi Mallick


Bruce Ray Wilson


Felix Perez

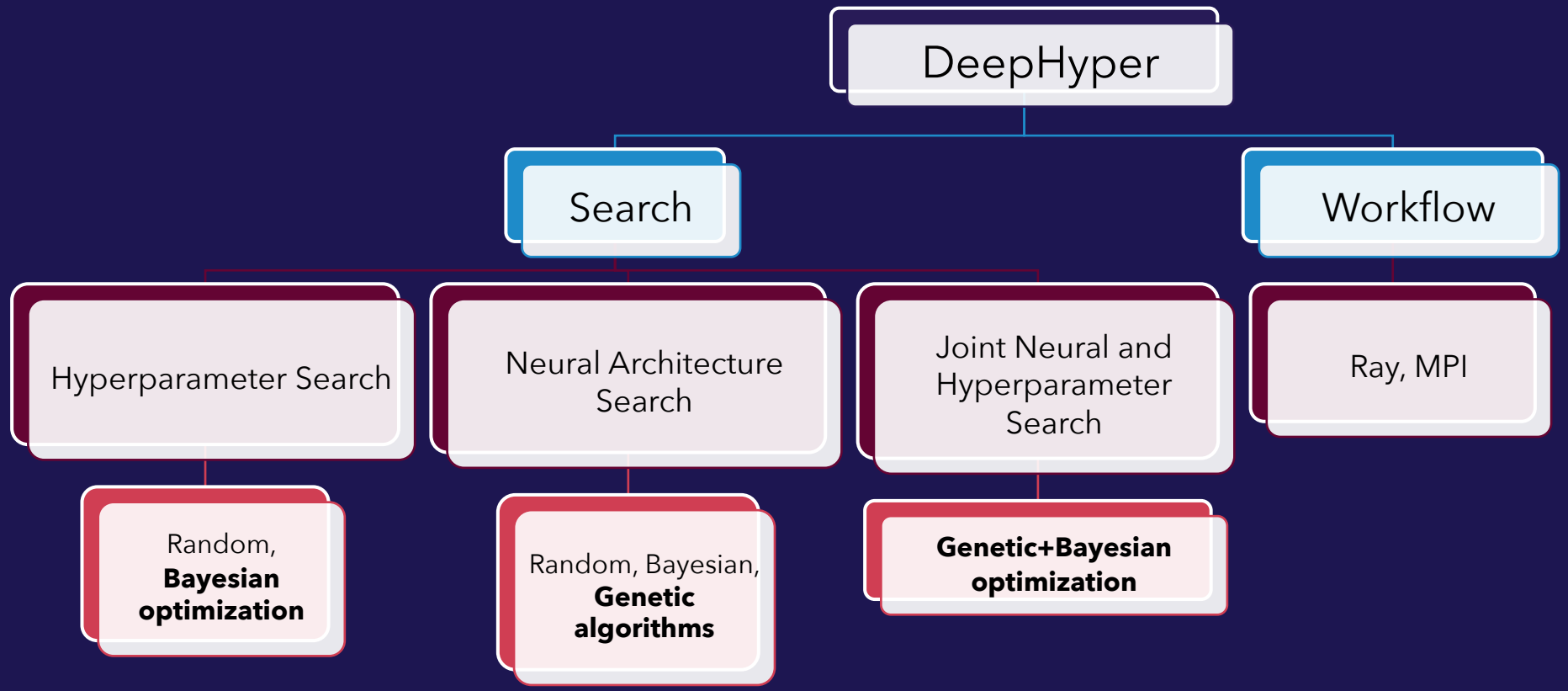
Acknowledgements

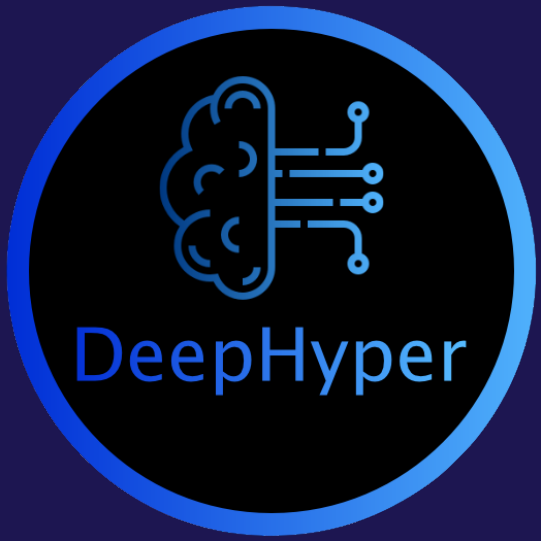
DOE Early Career Research Program, ASCR

Argonne Leadership Computing Facility

Laboratory Directed Research and Development (LDRD)

DeepHyper: Scalable AutoML





DeepHyper documentation: http://deephyper.readthedocs.io