FLORIDA TECH

FLORIDA'S STEM UNIVERSITY™

AUGMENTED POPULATION BASED TRAINING (APBT)

Josias Moukpe • 4-27-2022

OUTLINE

- Problem
- Approach
- Operators
- Results – Tennis
- Results – Iris
- Results – Identity
- Discussion – Advantages
- Discussion – Drawbacks
- Observations
- Summary

FLORIDA TECH

Problem

Searching and Tuning Neural Nets is a difficult and long process!

- Artificial Neural Nets (ANN) are very powerful function approximators
- To use effectively, ANNs require lots of hyperparameters to tune (6+, too much to exhaustively search!)
- Moreover, finding the right topology is even more important and difficult. Wrong topology invalidates best hyperparameters…

FLORIDA TECH

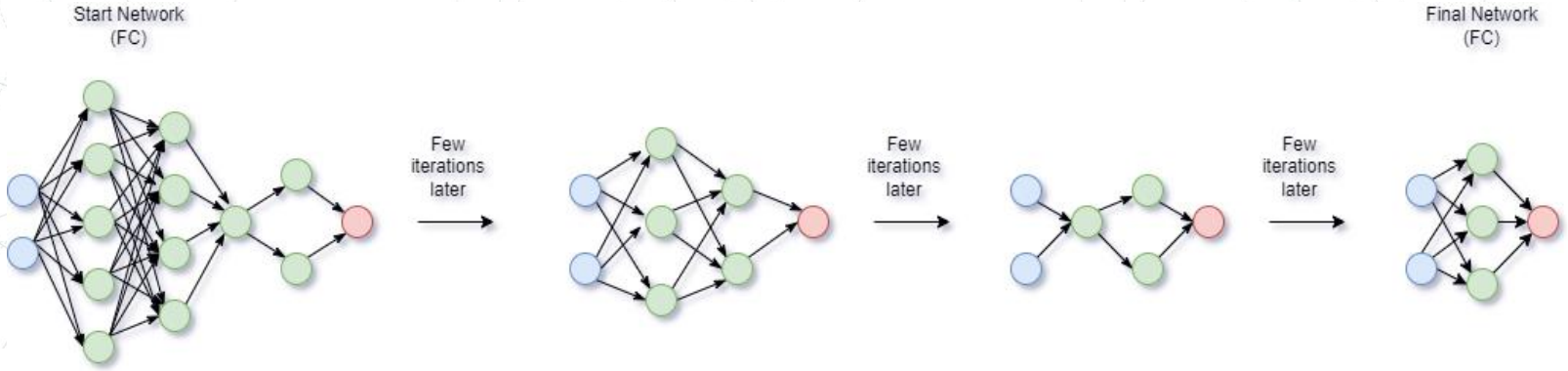
Approach

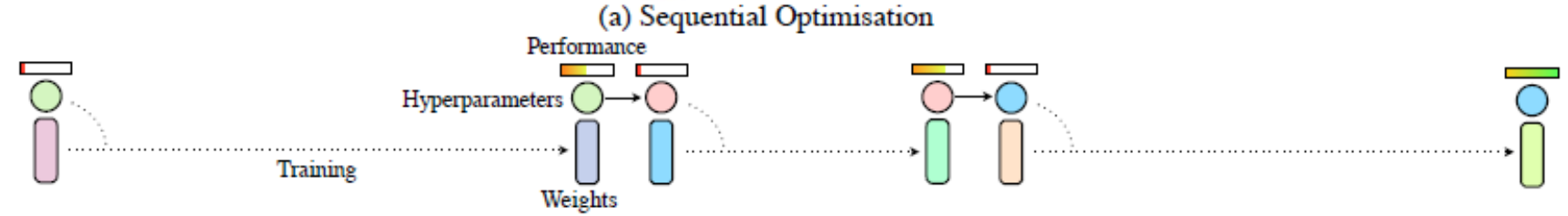
We propose Augmented Population Based Training!

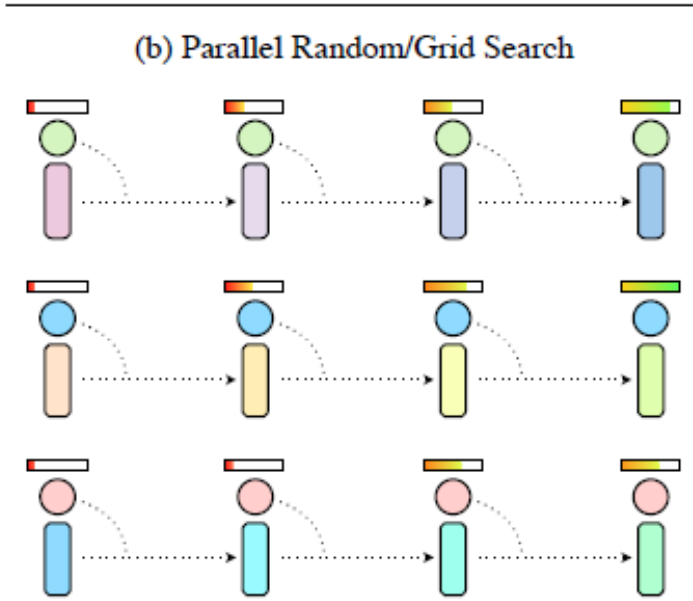
- Leverage power of Genetic Algorithms to help search both topology and hyperparameters in SINGLE joint training process.
- Produce optimal topology (high accuracy, low size) trained on optimal hyperparameter schedules (adaptive, not just one value)
- Requires fewer hyperparameters (3 at most, 2 usually!)
- Can be parallelize for speed boost.
- Better than manual, grid, and random search

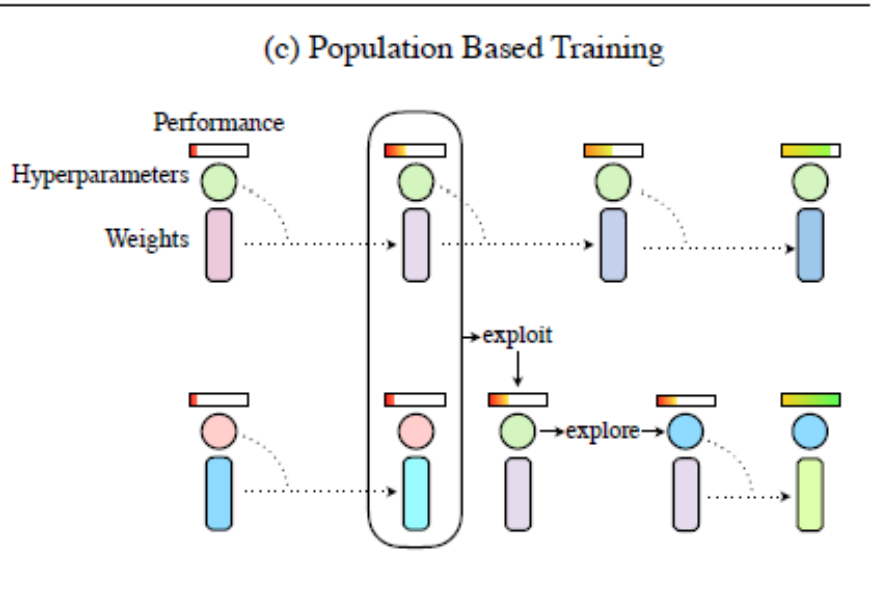
FLORIDA TECH

Illustrations







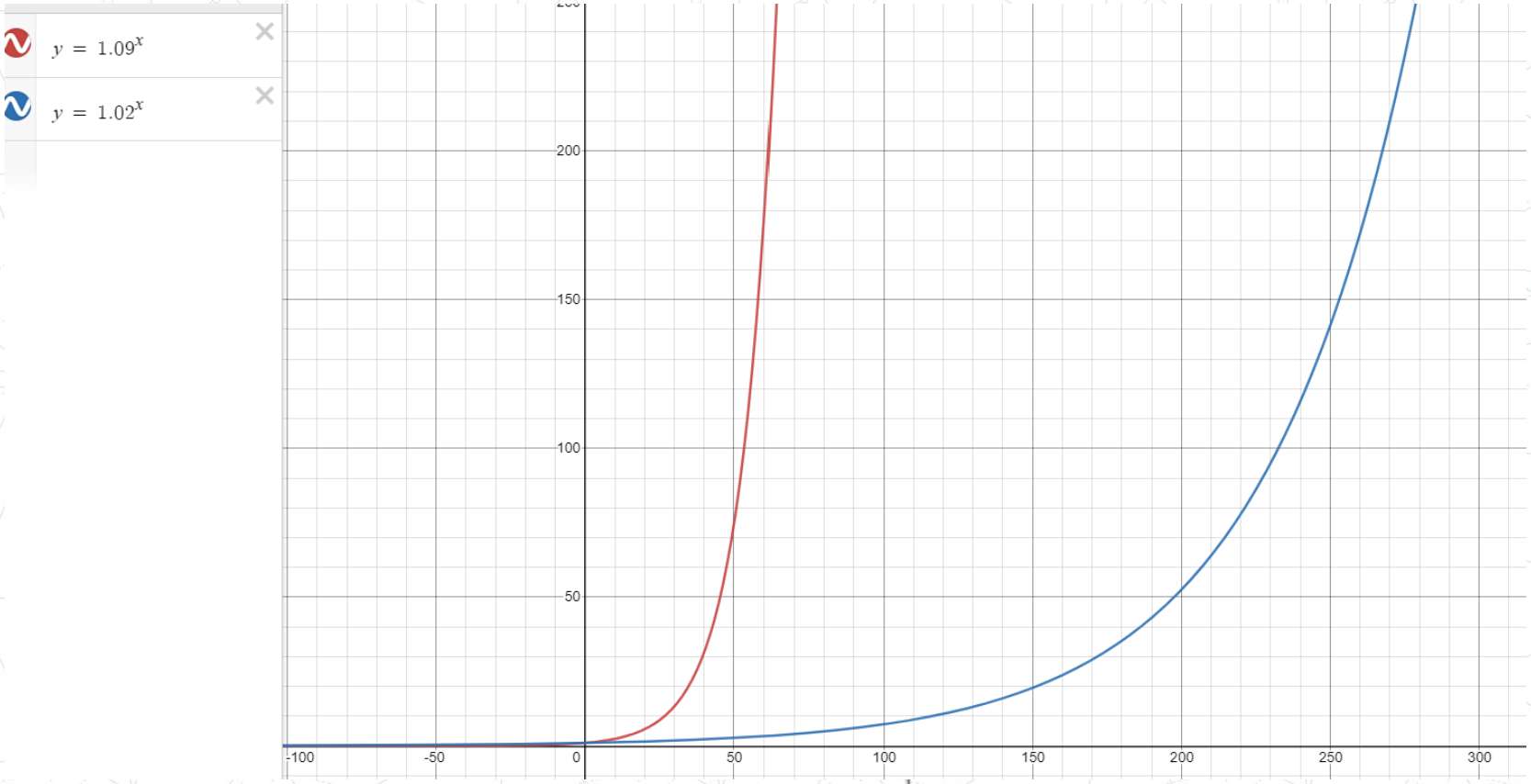


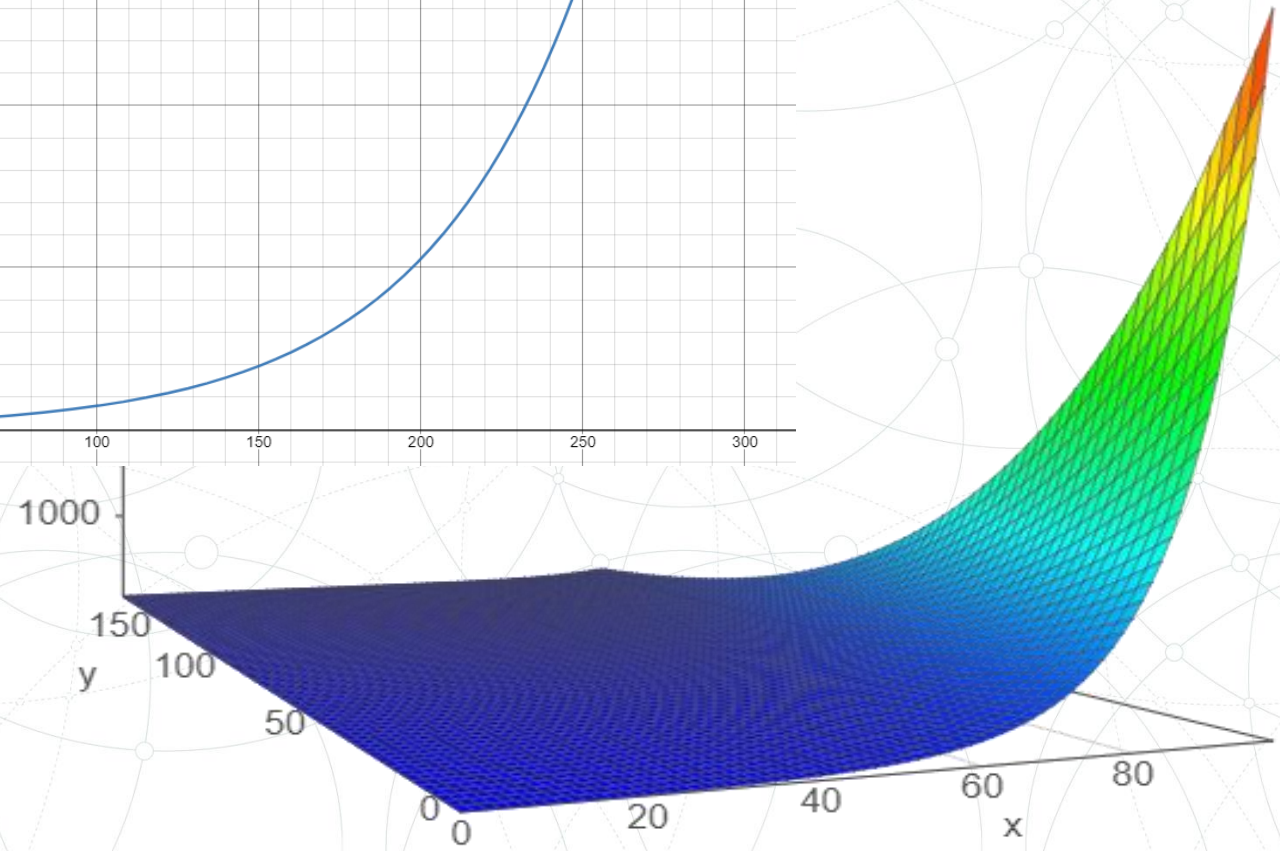
Augmented Population Based Training Operators

Evaluate

- It's our Fitness function here.
- Takes a neural net, a validation/test set, and returns a score
- Evaluate the accuracy (1), get net size (2), and apply a special ratio function (f) on the accuracy and size to get performance score.
- $f(accuracy, size) = \frac{1.09^{accuracy}}{1.02^{size}}$

FLORIDA TECH





Ratio F

FLORIDA TECH

Augmented Population Based Training Operators

Truncation Selection

- It's our selection function based on the readiness threshold.
- A neural net is simply "ready" after a set number of iteration has elapsed since last time it was ready
- Truncation selection works by first ranking the nets in the population by performance ratio. If current net is in the bottom 20% of the population, we sample another net uniformly from the top 20% of the population to exploit.

FLORIDA TECH

Augmented Population Based Training Operators

Exploit

- It’s our crossover function.
- When a net is ready after a certain number of iterations, and is struggling in performance (bottom 20%), it can exploit the rest of the population by copying the topology, weights, and hyperparameters of the top net selected from the population with truncation selection.

FLORIDA TECH

Augmented Population Based Training Operators

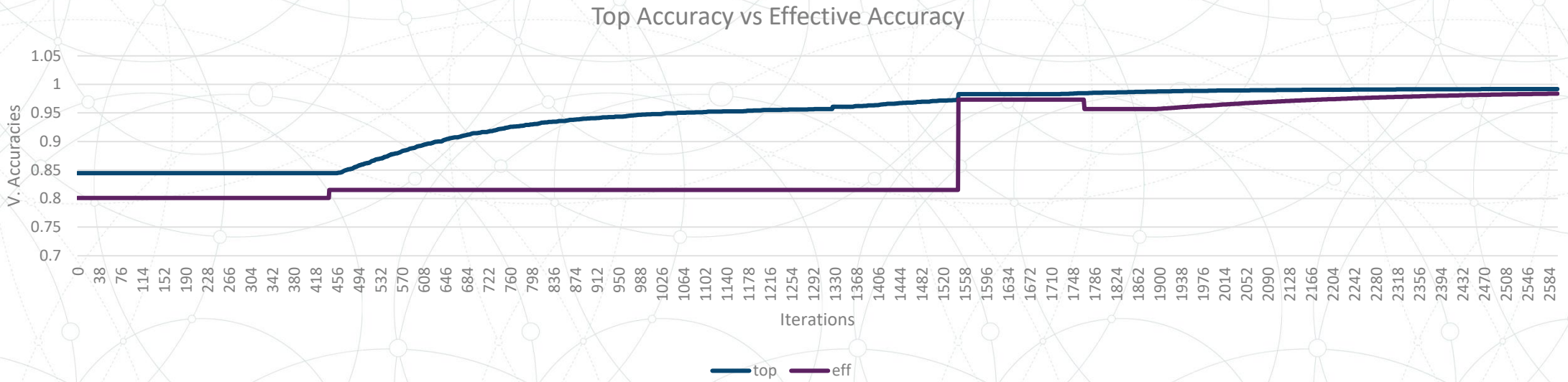
Explore

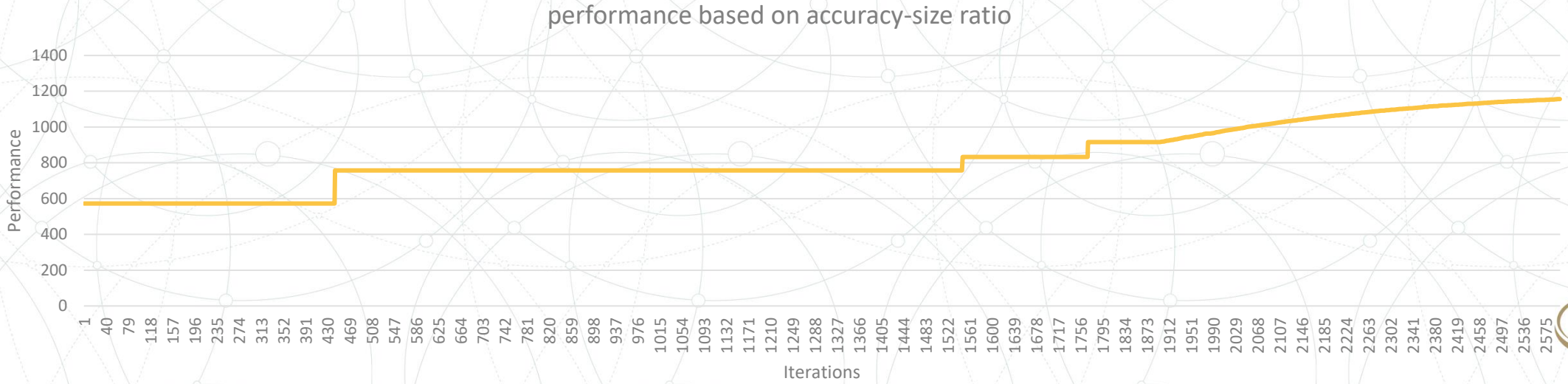
- It’s our mutation function.
- Following exploitation, the net experiences a perturbation in its hyperparameters, and topology before continuing training with the new hypermeters and topology.
- Hyperparameters are multiplied by either .8 or 1.2
- Topology either gains a neuron, loses a neuron, or does nothing.

FLORIDA TECH

Results

Tennis Dataset

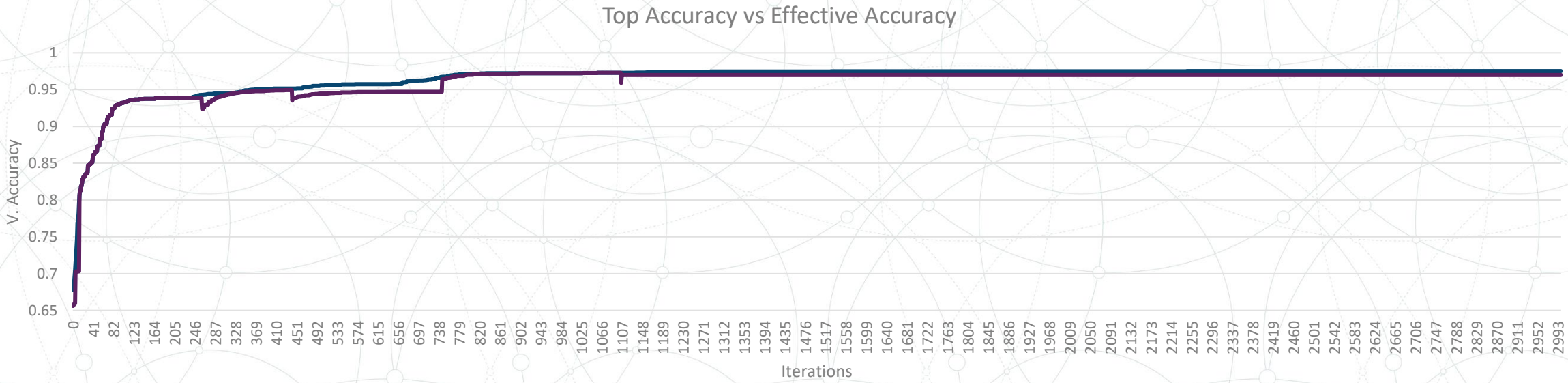


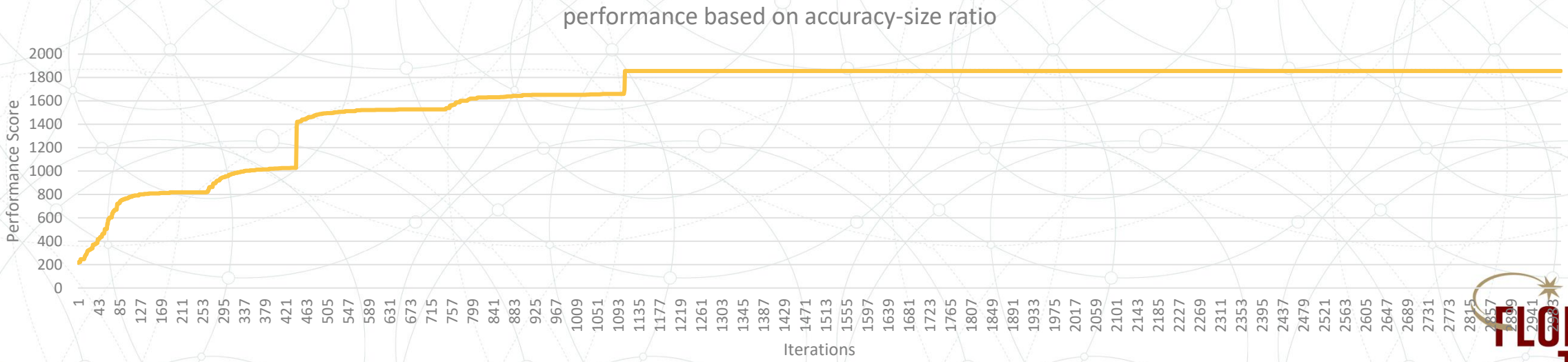


FLORIDA TECH

Results

Iris Dataset

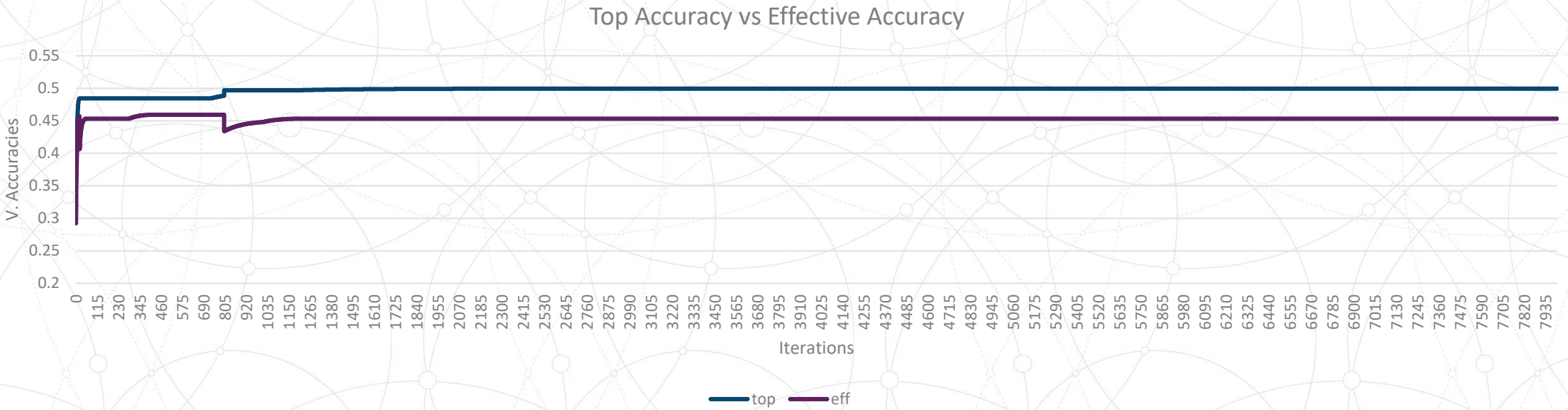


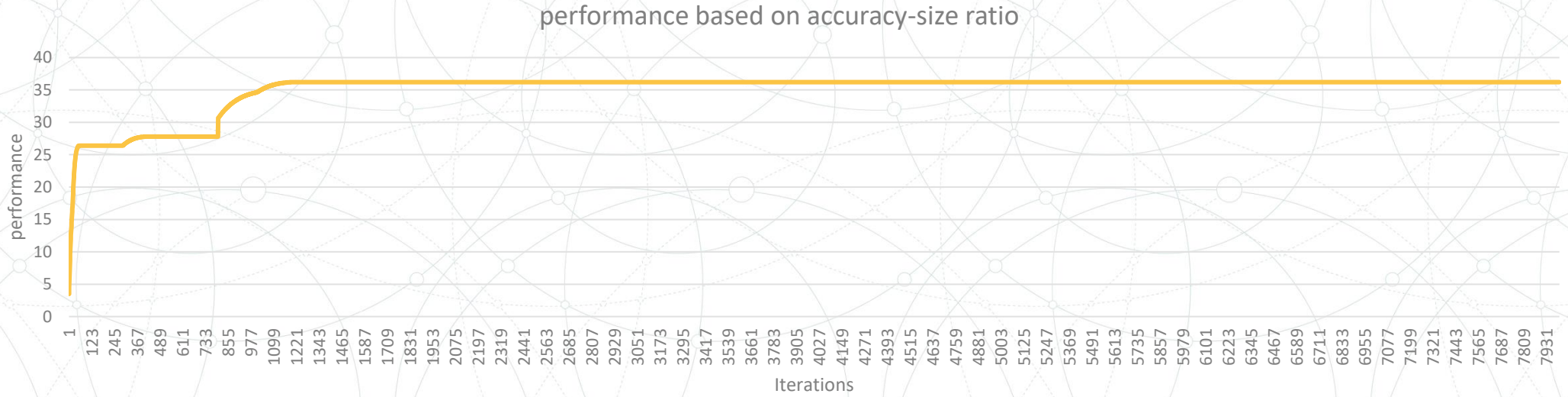


FLORIDA TECH

Results

Identity Dataset





FLORIDA TECH

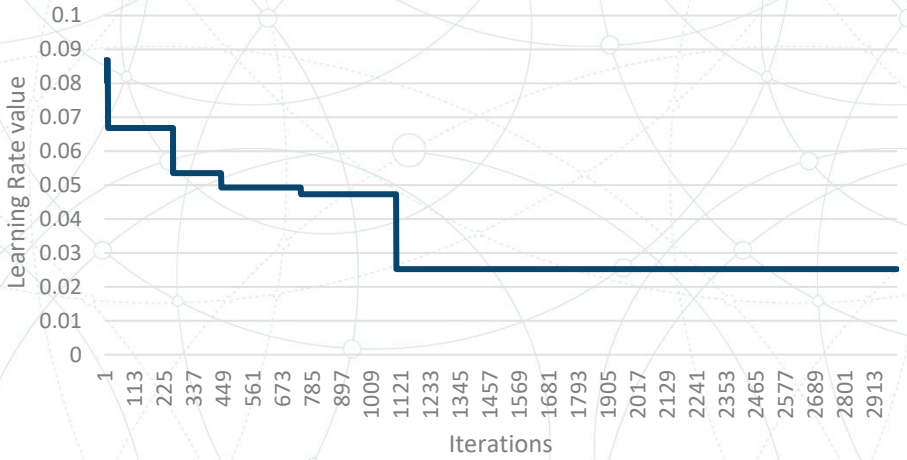
Results

	Iris					
	Trial 0		Trial 1		Trial 2	
	Best	Most acc	Best	Most acc	Best	Most acc
Test Acc %	***94.24***	***96.6***	***91.24***	***94.88***	***66.42***	***68.8***
Num Params	***42***	***63***	***28***	***49***	***7***	***70***
Topology	***4,6,3***	***4,9,3***	***4,4,3***	***4,7,3***	***4,1,3***	***4,10,3***
	Tennis					
	Trial 0		Trial 1		Trial 2	
	Best	Most acc	Best	Most acc	Best	Most acc
Test Acc %	***96.72***	***97.14***	***85.33***	***90.59***	***92.49***	***95.77***
Num Params	***72***	***120***	***12***	***96***	***12***	***24***
Topology	***10,6,2***	***10,10,2***	***10,1,2***	***10,8,2***	***10,1,2***	***10,2,2***
	Identity	(struggles here…)				
	Trial 0		Trial 1		Trial 2	
	Best	Most acc	Best	Most acc	Best	Most acc
Test Acc %	***55.2***	***93.38***	***54.81***	***92.39***	***55.32***	***89.11***
Num Params	***16***	***144***	***16***	***160***	***16***	***144***
Topology	***8,1,8***	***8,9,8***	***10,1,2***	***8,10,8***	***8,1,8***	***8,9,8***

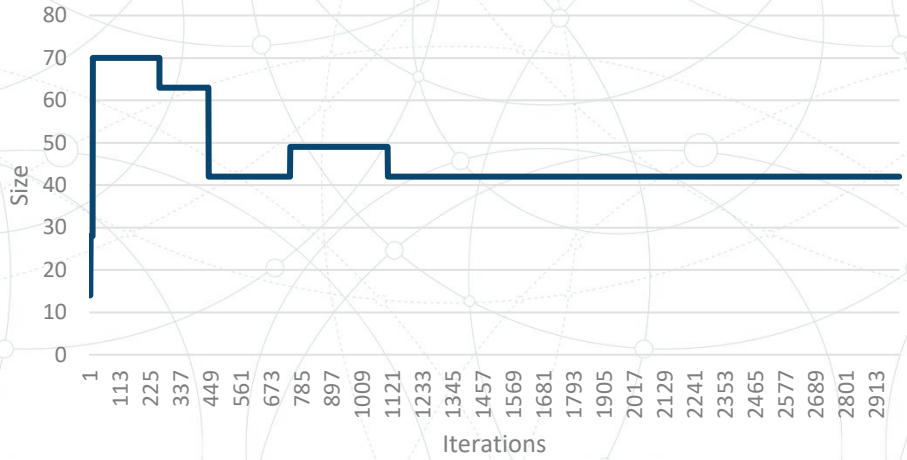
FLORIDA TECH

Hyperparameter schedules (think of what they mean)

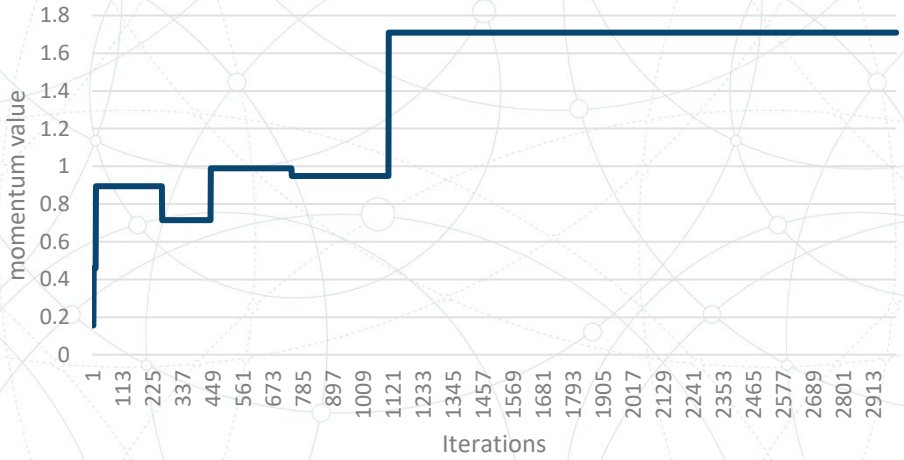
adaptive learning rate schedule



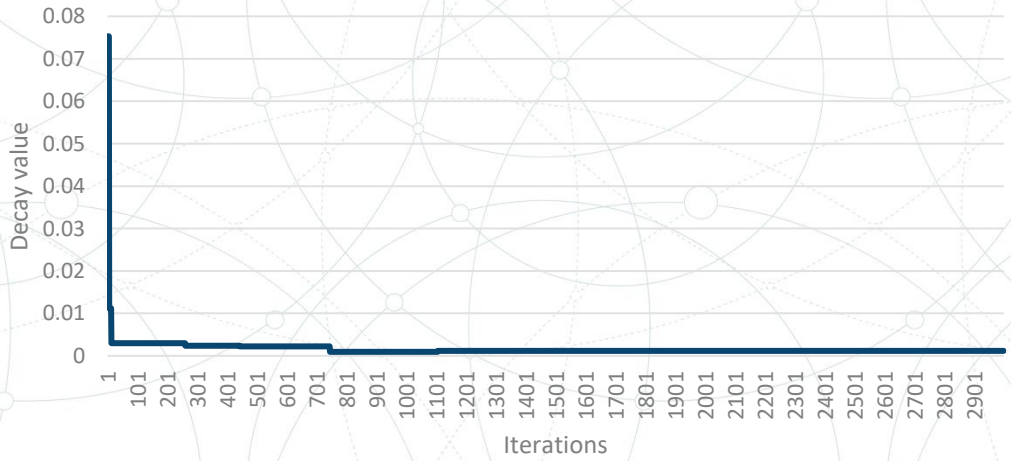
Size or Number of learnable parameters



momentum rate schedule



weight decay schedule



FLORIDA TECH

Discussion: Advantages

- Works quite well in optimizing topology and hyperparameters!
- Provides both absolute highest accuracy net and highest accuracy with lowest size net.
- Trains ANN adaptively on hyperparameter schedules as opposed to values for efficient training
- Requires much less tunable parameters.
- Found 10.2.2 at 96% on Tennis
- Not sensitive to local minima or overfitting
- Can be parallelized

FLORIDA TECH

Discussion: Drawbacks

- APBT uses **k** (population size) times more space to run than BP, runs **k** times longer. But it can be parallelized to run much faster!
- APBT is inconsistent with too small **k** and to small datasets
- APBT struggles to find minimum size net with Identity dataset
- Doesn’t support early stopping yet

FLORIDA TECH

Some Observations

- Higher k the better, as long as resources allow it. Increases diversity in population
- Higher Selection Readiness allows nets to train for longer before sharing. But too high readiness and it's just k-random search.
- Most effectively accurate net is not always the most absolutely accurate
- Longer epochs range allow for rest of population to catch up to the top performers.

FLORIDA TECH

Summary

- APBT can work really well and requires far less hyper parameters, providing the most accurate, lowest size neural net, which is trained on the optimal hyperparameter schedule.
- However, APBT takes k (population size) times more space to run than BP, and without parallelism runs k times longer than BP. Can be inconsistent with smaller k and very small datasets. Struggles with Identity
- In all, to solve the problem of ANN topology and hyperparameters optimization, Augmented Population Based Training is a powerful step in the right direction.

FLORIDA TECH

FLORIDA TECH

FLORIDA'S STEM UNIVERSITY™

Thank you.

Josias Moukpe • 4-27-2022


Questions ?