

Auto-tuning of Hyperparameters of Machine Learning Models

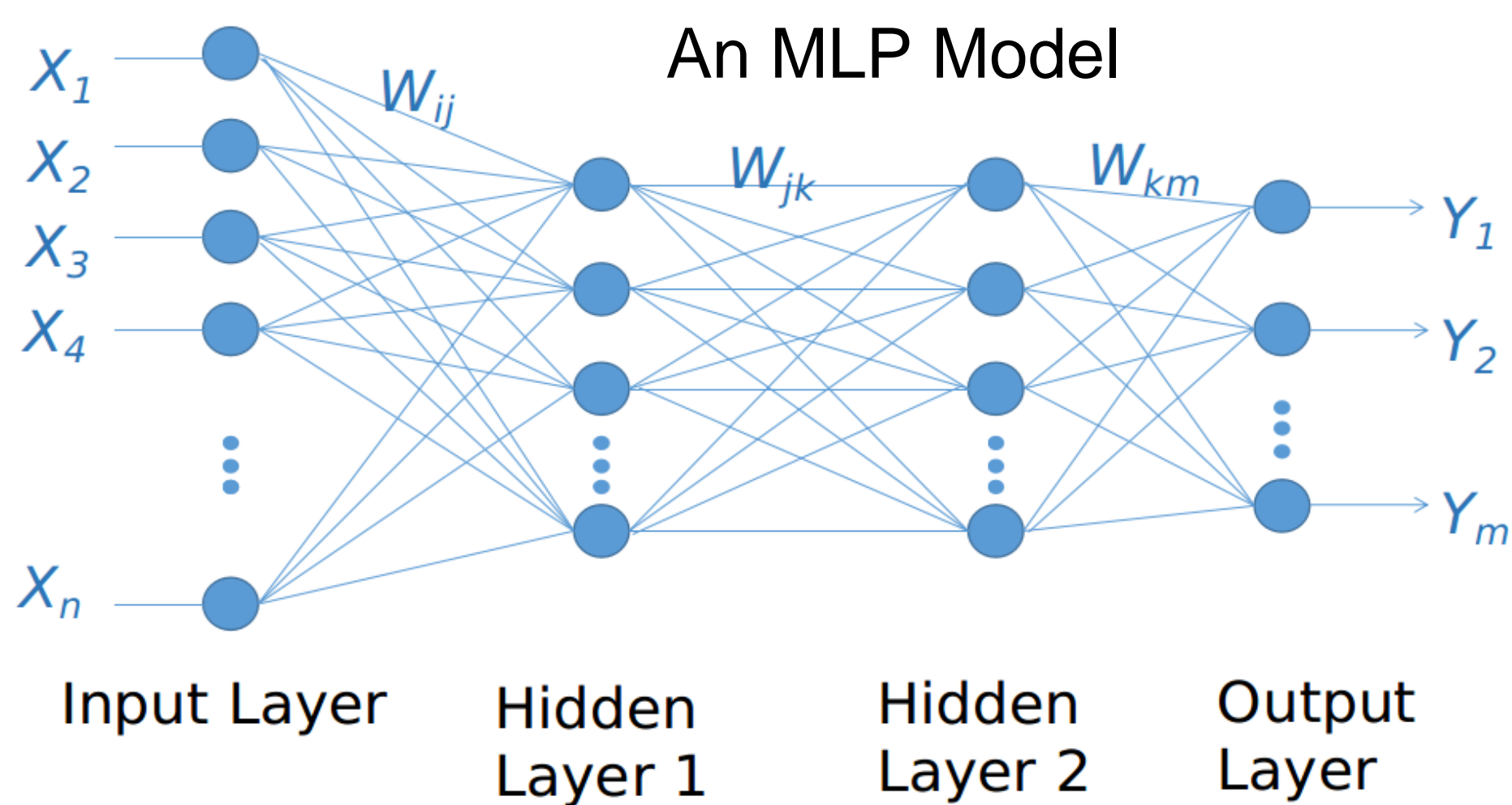
Zhen Wang¹, Ryusuke Egawa¹, Reiji Suda², Hiroyuki Takizawa¹,
¹: Tohoku University; ²: The University of Tokyo

Introduction

Background and Motivation

Parameters & Hyperparameters:

Suppose a simple **Machine Learning model** – a four-layer multilayer perceptron (MLP):



A typical training process would be:

- Configure **hyperparameters** (the number of hidden layers etc.)
- Training process starts, standard **parameters** will be founded

Accordingly, two kinds of parameters, hyperparameters and (standard) parameters need to be properly configured.

How to find a good configuration of hyperparameters?

Two common methods to find a good configuration

- **Full(Grid) search**: a hyperparameter space is discretized, and an MLP model is trained and assessed for all combinations of the values over the space, which is too costly.

- **Random Combination**: built with a random combination of hyperparameter values. The efficiency strongly depends on random combinations.

As the performance of a trained machine learning model usually has an uncertainty, we need to assess the performance of one hyperparameter configuration N times (N trials) to obtain the average performance. Thus, if there are M parameter configurations, one simple way is to repeat training and assessing a model $N \times M$ times, i.e., **$N \times M$ trials**, to decide the best parameter configuration in terms of the average performance.

Can we auto-tune the hyperparameter configuration selection with less trails?

Proposed Approach/Experiment setup/Evaluation

ATMathCoreLib[1] is an **auto-tuning mechanism** originally developed for improving the performance of a program by adjusting its parameters affecting the execution time in the **HPC field**. It can adjust a lot of parameters based on the Bayesian optimization under the situation where the performance of a model defined by the parameters has an uncertainty, i.e., the observed performance may have some perturbations.

Thus, in our work, we focus on the **feasibility** of using such technologies to select the configuration hyperparameters of a machine learning model with less trials. The typical usage of ATMathCoreLib for finding a good configuration of hyperparameters is:

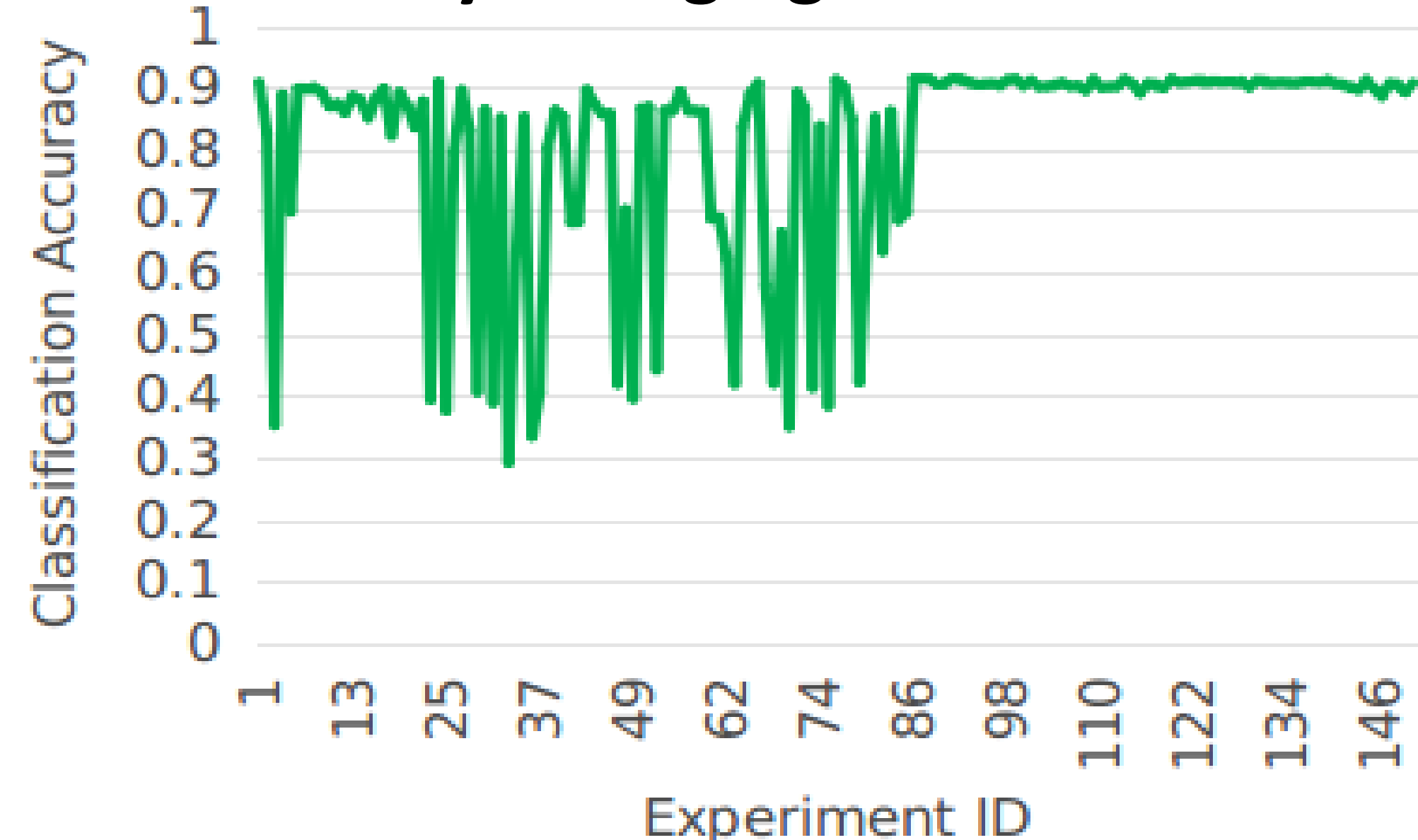
- Machine Learning models selection
- Hyperparameters selection with ATMathCoreLib
- Training process starts

In the experiment, two hyperparameters: the hidden layers and the number of neurons in each hidden layer of an MLP model are adjusted using ATMathCoreLib. The experiment setup is:

Machine Learning Model	Data Sets	Experiment Environment
Multi-layer Perceptron	MNIST[2]	i7-3770 CPU, 3.4GHZ

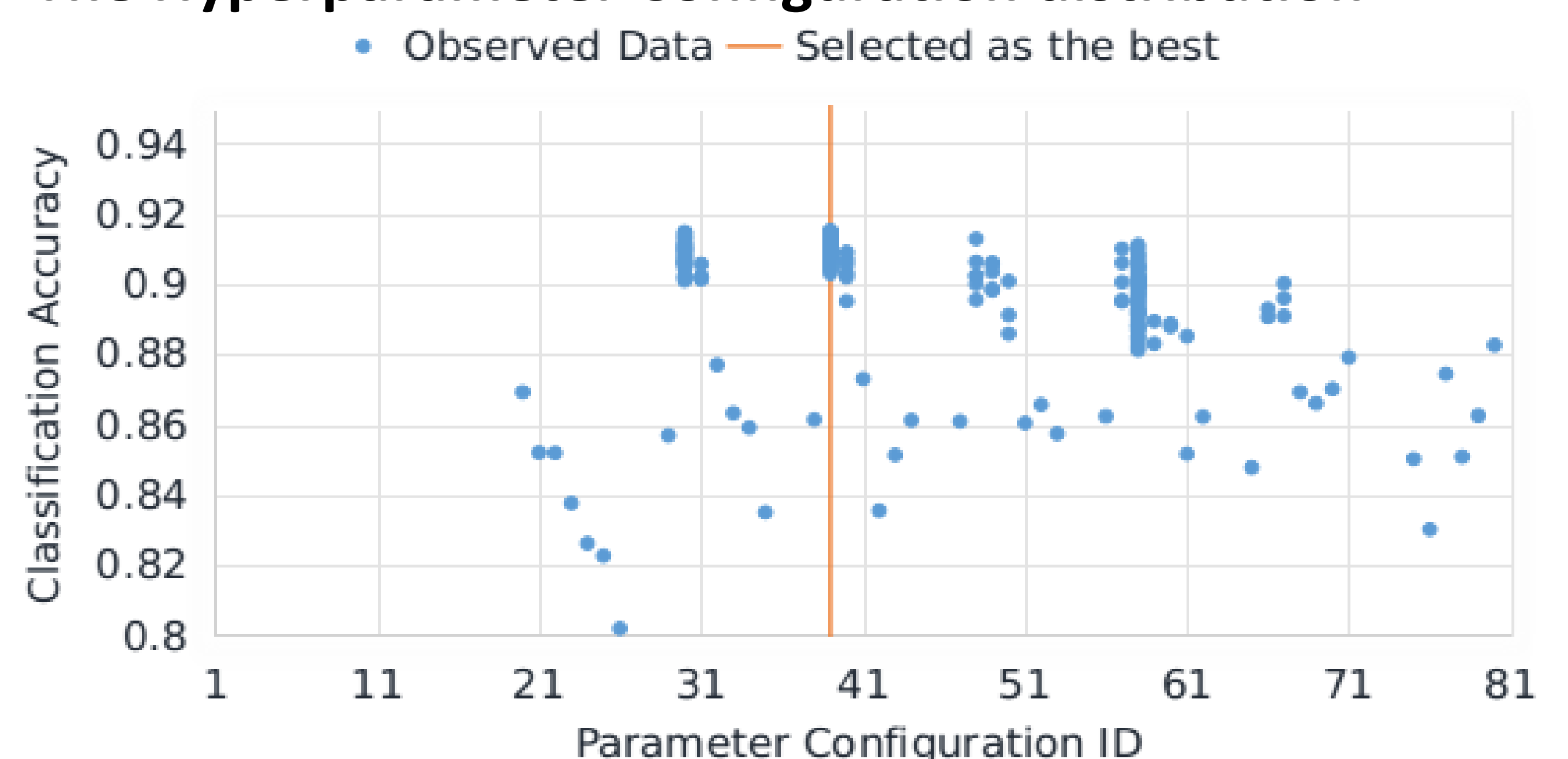
A unique ID named experimental ID is given to each trial and the parameter configuration ID is given to each configuration of hyperparameters.

The Accuracy Changing Trend



After about 100 trials in this experiment, only hyperparameter configurations with **higher classification accuracies** are identified as promising ones and tested to find the best one among them.

The Hyperparameter Configuration distribution



In the case of using ATMathCoreLib, hyperparameters with **higher classification accuracies** are likely tested more times than the others. Our experiments show that ATMathCoreLib can efficiently find an appropriate parameter configuration with less trials.

Conclusions and Future Work

This poster shows an auto-tuning mechanism for the configuration of hyperparameters of Machine Learning models. The evaluation results show that **auto-tuning** technologies in the HPC field are also **effective** to intelligently adjust the **hyperparameters** of machine learning models. In our future work, we will use more advanced machine learning models and discuss the effectiveness of auto-tuning in more practical situations.

[1] Reiji Suda. 2011. A Bayesian method of online automatic tuning. In Software Automatic Tuning. Springer, 275–293

[2] Cortes Corinna LeCun Yann. 2017. The MNIST Database of Handwritten Digits. (2017). <http://yann.lecun.com/exdb/mnist/>