APRIL 2024

Energy-Accuracy Trade off in Hyperparameter Tuning

A study on how hyperparameter tuning affects the accuracy and energy consumption of machine learning models.

Abdulrahman Alshahrani, Abdallah Al-Sukhni
Department of Computer Science, The University of Texas at Austin

Agenda:

1. Energy consumption of Machine Learning
2. What is Hyperparameters Tuning?
3. The issue with Grid Search
4. Our Experiment
5. Our Results
6. Conclusion

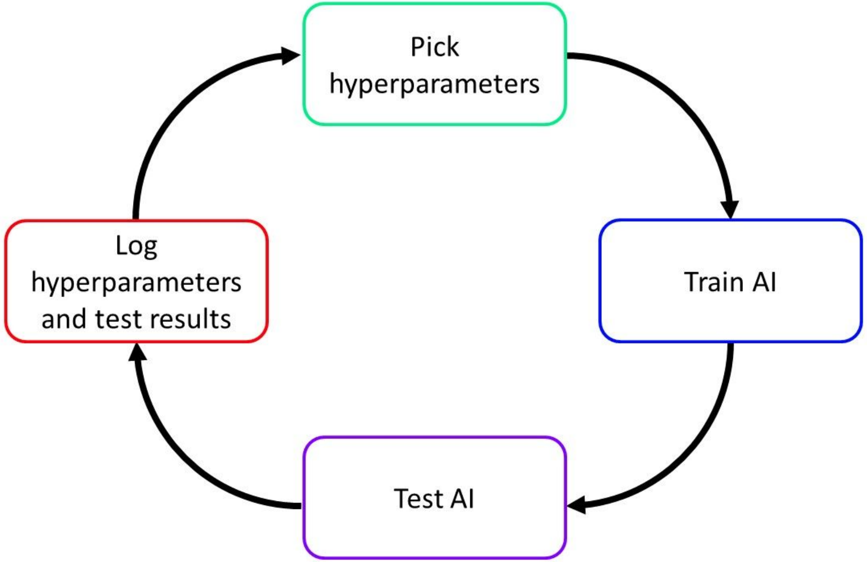
Machine Learning Energy Consumption

- AI could consume **85-134** TWh by 2027 [1]
- Competes with **Netherlands** annual energy consumption. [1]
- Training GPT-3 emitted **502** metric tons of carbon, equivalent to driving **112** gasoline powered cars for a year. [2]

[1] How much electricity does AI consume?
Vincent
https://www.theverge.com/24066646/ai-electricity-energy-watts-generative-consumption

[2] https://news.climate.columbia.edu/2023/06/09/ais-growing-carbon-footprint/

How does an ML Model Works



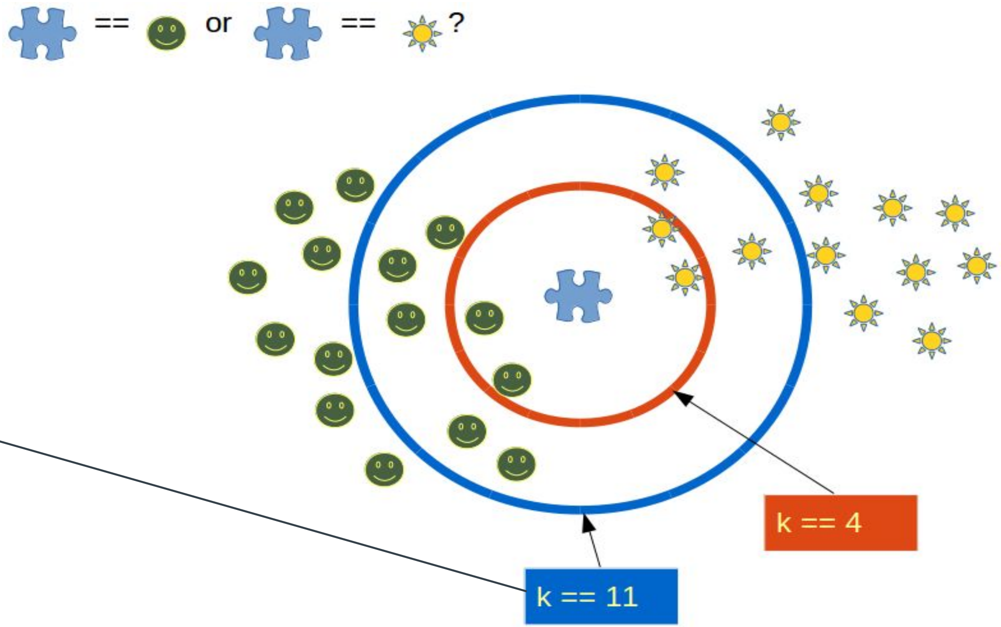
But What’s a Hyperparameter?

- It specifies the **structure** of your ML model.
- You set it **before** training your model.

But What's a Hyperparameter?

E.g.

- K-nearest neighbors



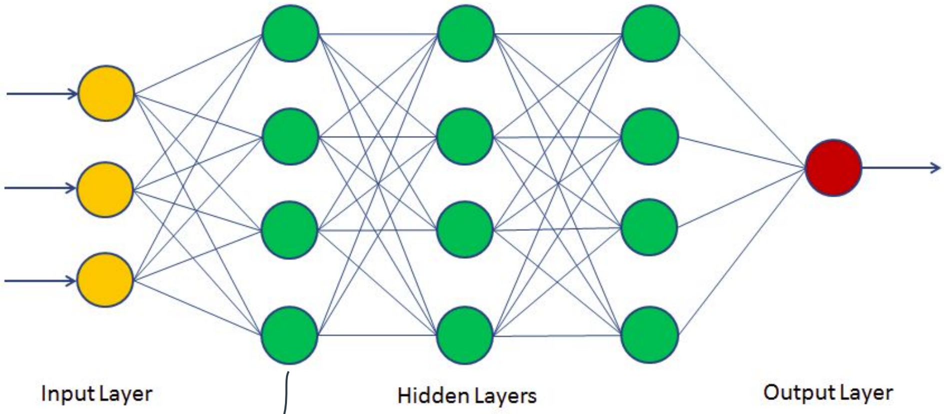
What K value to use?

https://towardsdatascience.com/model-selection-tuning-and-evaluation-in-k-nearest-neighbors-6d3024d78745

But What’s a Hyperparameter?

E.g.

- Neural Networks



https://ai.plainenglish.io/introduction-to-artificial-neural-network-c1aaf4929c03

How many nodes in each layer?

how many hidden layers?

How to choose the **optimal** Hyperparameters?

(aka Hyperparameter Tuning)

- First… What is “*optimal*”?
- Optimal = the highest accuracy for the model

How to choose the **optimal** Hyperparameters?

(aka Hyperparameter Tuning)

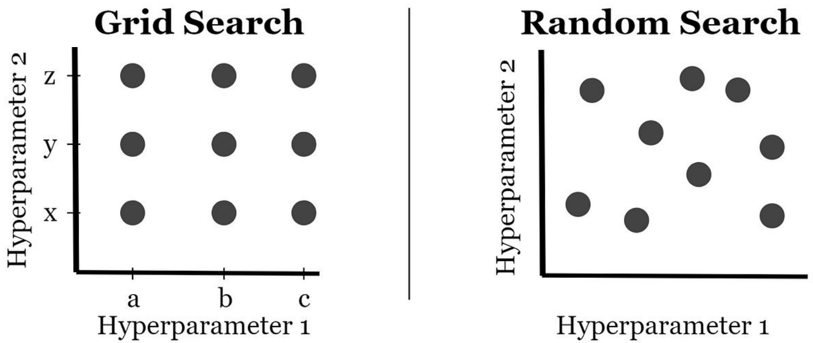
- Background experience (too hard, not reliable)

Or…

- Let the machine try many values

What values to try?

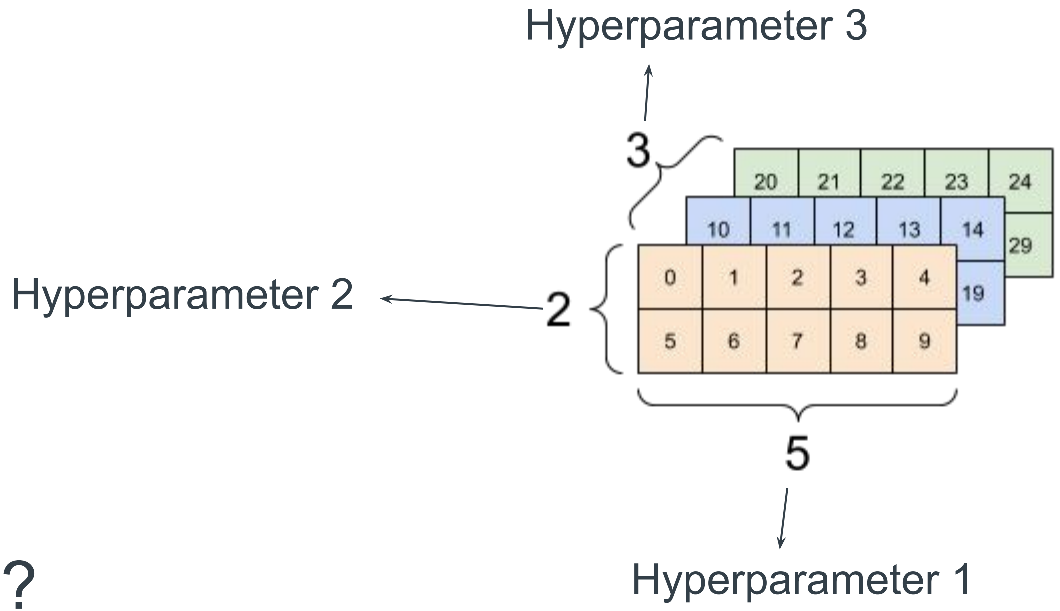
- **Testing all values**
    - **aka Grid Search**
- Test Random values
    - aka Randomized Search



The issue with Grid Search

- Exponential

- Maybe inefficient?



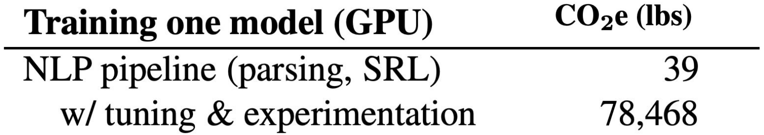
Hyperparameter 3

Hyperparameter 2

Hyperparameter 1

The issue with Grid Search

- **~2000x** increase in CO_2 [1]



Training one model (GPU)	CO_2e (lbs)
NLP pipeline (parsing, SRL)	39
w/ tuning & experimentation	78,468

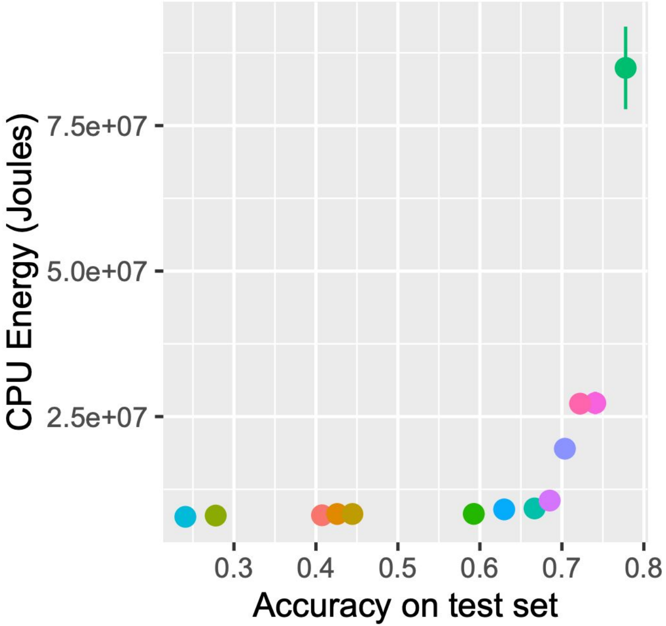
- **77%** decrease in energy consumption vs. **1%** decrease in accuracy [2]

[1] E. Strubell, A. Ganesh, and A. Mccallum, “Energy and Policy Considerations for Deep Learning in NLP,” 2019.
[2] A. Brownlee, J. Adair, S. Haraldsson, and J. Jabbo, “Exploring the Accuracy -Energy Trade-off in Machine Learning.”

So Who Cares?

What Others Did

- Brownlee, Adair, Haraldsson, Jabbo
- They investigated the energy and accuracy trade-offs of using grid search on multiple hyperparameter options across several datasets
- Findings revealed a clear opportunity to trade off small accuracy in exchange for a much higher reduction in energy consumption

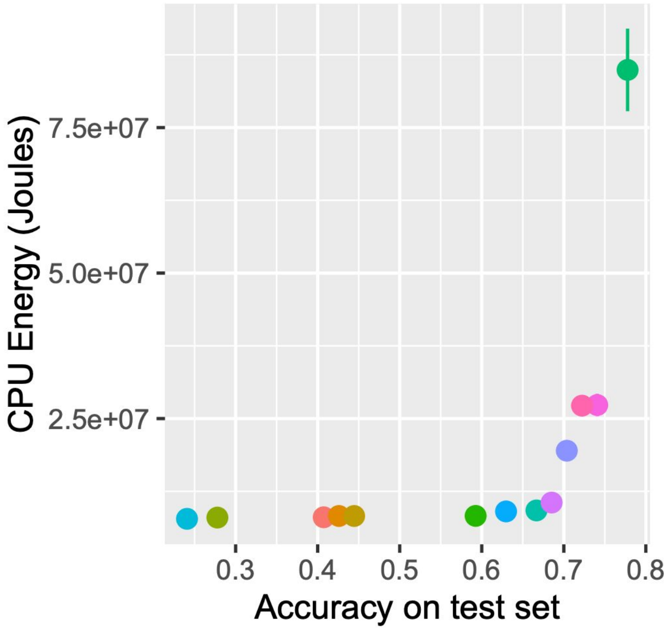


[1] A. Brownlee, J. Adair, S. Haraldsson, and J. Jabbo, “Exploring the Accuracy -Energy Trade-off in Machine Learning.”

What Others Did

- Energy savings of **~40%** with minimal reductions in classification accuracy.

- They only focused on MultiLayer Perceptrons (MLP)



[1] A. Brownlee, J. Adair, S. Haraldsson, and J. Jabbo, “Exploring the Accuracy -Energy Trade-off in Machine Learning.”

What We’re Doing

- Can we get similar savings with other algorithms?
    - Multilayer Perceptrons (MLPs)
    - K Nearest Neighbors
    - Decision Trees
    - Random Forest

Our Setup

- The Algorithms
- The Datasets
- Hyperparameter Grids
- The Machines
- Measuring Energy Consumption

The Algorithms

- Grid Search
  - GridSearchCV()
- Classifiers
  - DecisionTreeClassifier()
  - KNeighborsClassifier()
  - MLPClassifier()
  - RandomForestClassifier()

The Datasets

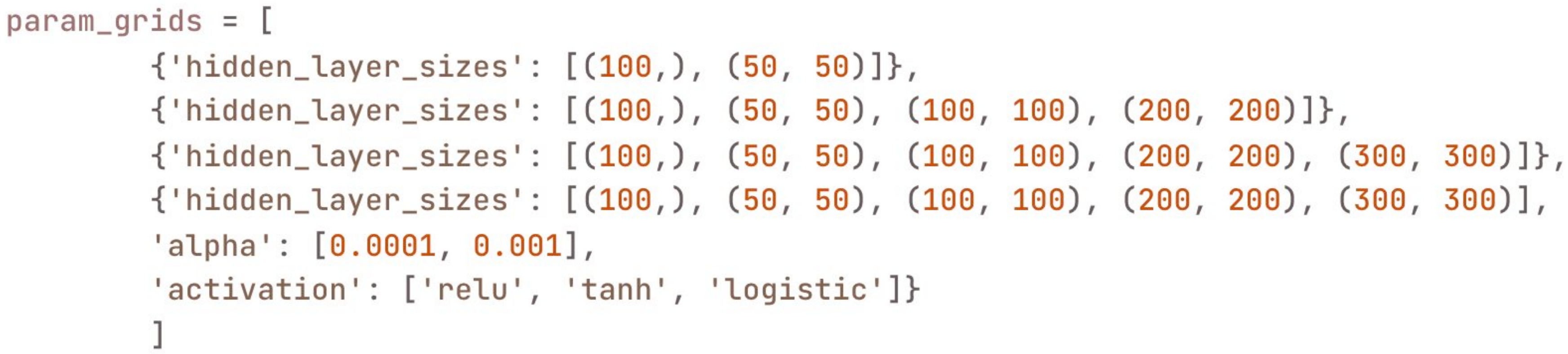
- Olivetti Faces (from scikit-learn)
  - ~400 instances
- Census Income
  - ~50k instances
- CDC Diabetes Health Indicators
  - ~ 250k instances
- RT-IoT2022
  - ~120k instances

scikit learn

UCI Machine Learning Repository

Hyperparameter Grids

For example, MLP’s:



```
param_grids = [
        {'hidden_layer_sizes': [(100,), (50, 50)]},
        {'hidden_layer_sizes': [(100,), (50, 50), (100, 100), (200, 200)]},
        {'hidden_layer_sizes': [(100,), (50, 50), (100, 100), (200, 200), (300, 300)]},
        {'hidden_layer_sizes': [(100,), (50, 50), (100, 100), (200, 200), (300, 300)],
        'alpha': [0.0001, 0.001],
        'activation': ['relu', 'tanh', 'logistic']}
        ]
```

The Machines

- Abdulrahman’s Machine:
  - **CPU:** Intel Core i7-9750H, 6 cores, (Base: 2.60GHz, Max: 4.50 GHz)
  - **GPU:** NVIDIA GeForce GTX 1660 Ti
  - **Memory:** 32 GB
- Abdallah’s Machine:
  - **CPU:** Intel Core i5-9400, 6 cores, (Base: 2.90GHz, Max: 4.10 GHz)
  - **GPU:** NVIDIA GeForce GTX 1660 Ti
  - **Memory:** 16 GB

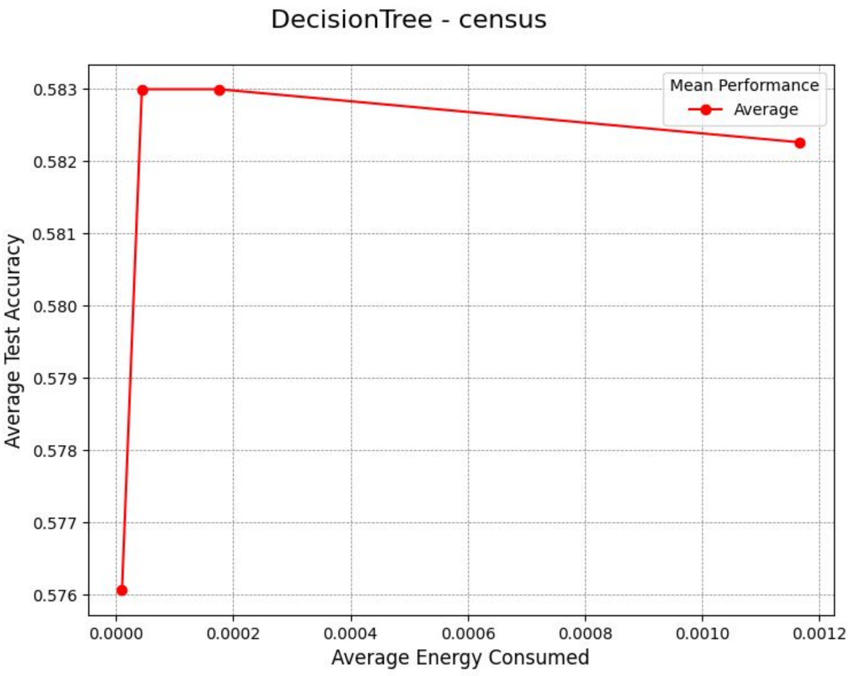
Measuring Energy Consumption

- **CodeCarbon** Python Library
  - Open Source
  - Lightweight
- Polling components for power consumption every some interval
  - We used every 5 seconds

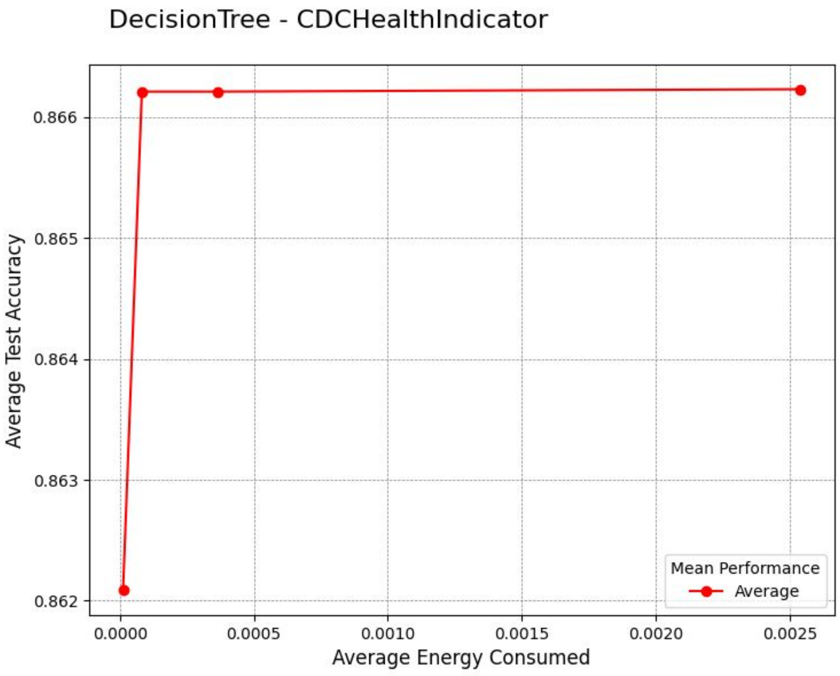
Measuring Energy Consumption

- GPU Power
  - Using **NVIDIA Management Library** on pyNVML
- CPU Power
  - Using **Intel Power Gadget**

The Results - Decision Trees

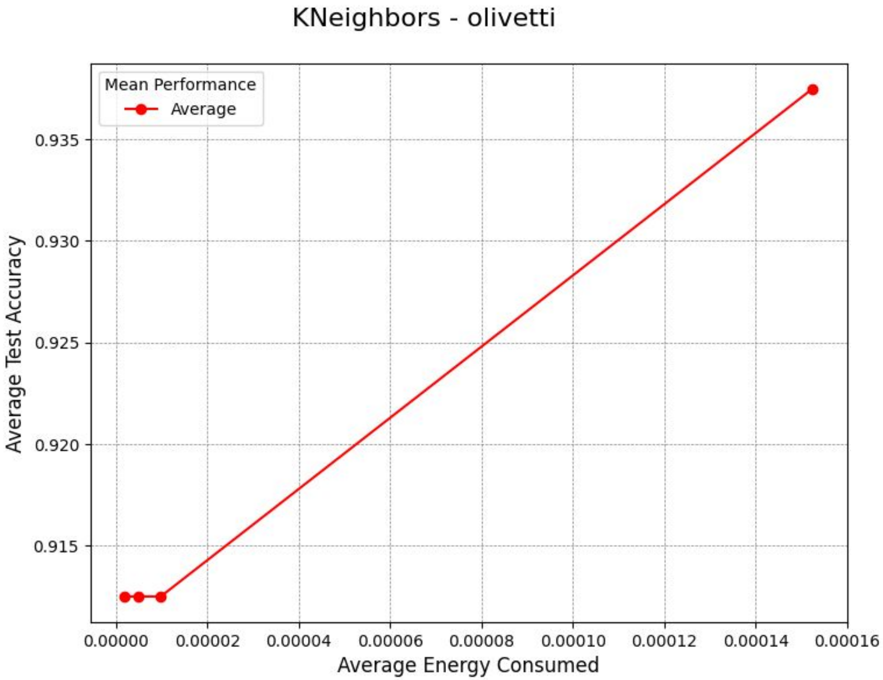


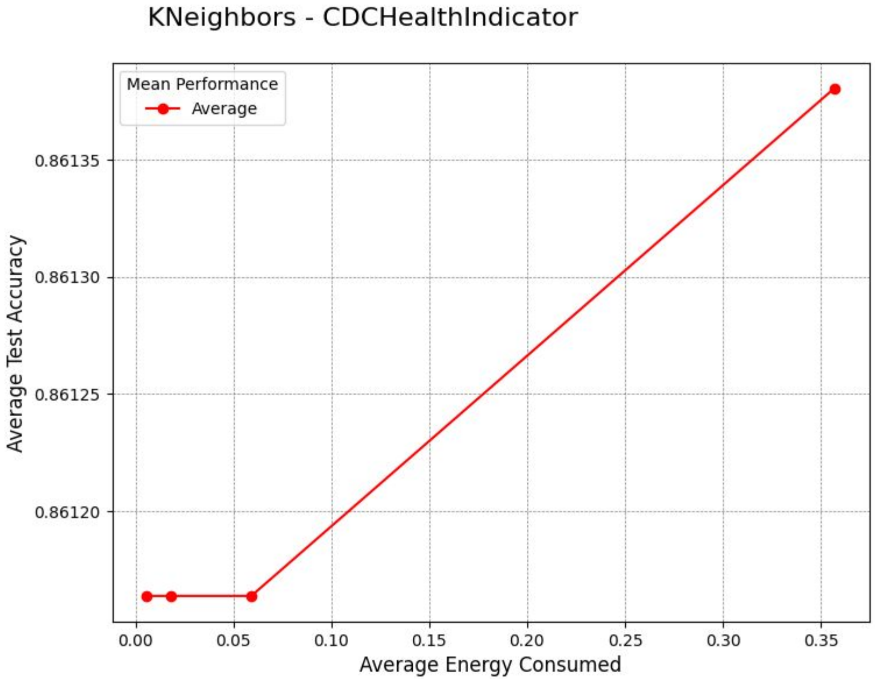
DecisionTree - census



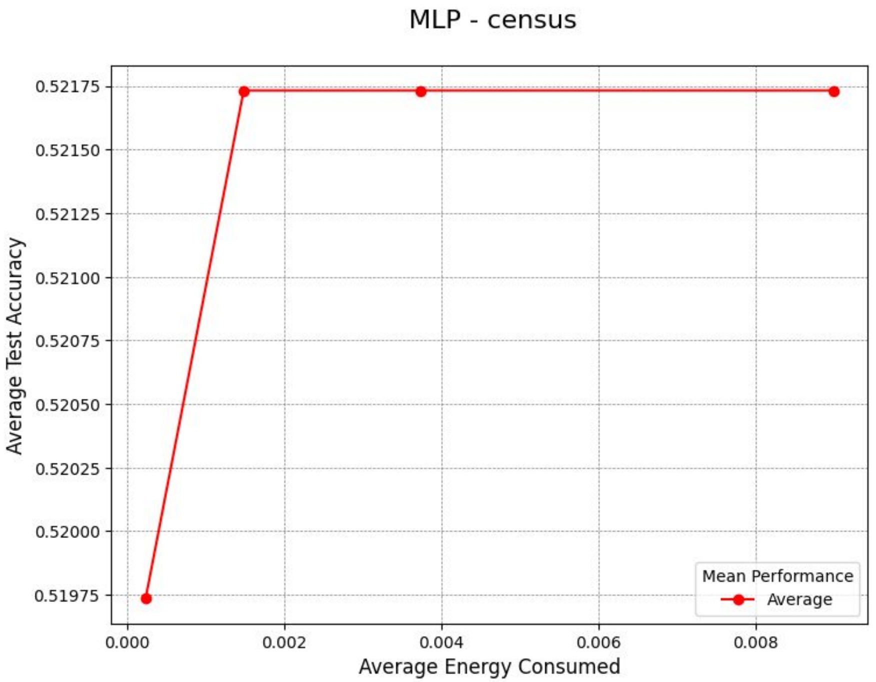
DecisionTree - CDCHealthIndicator

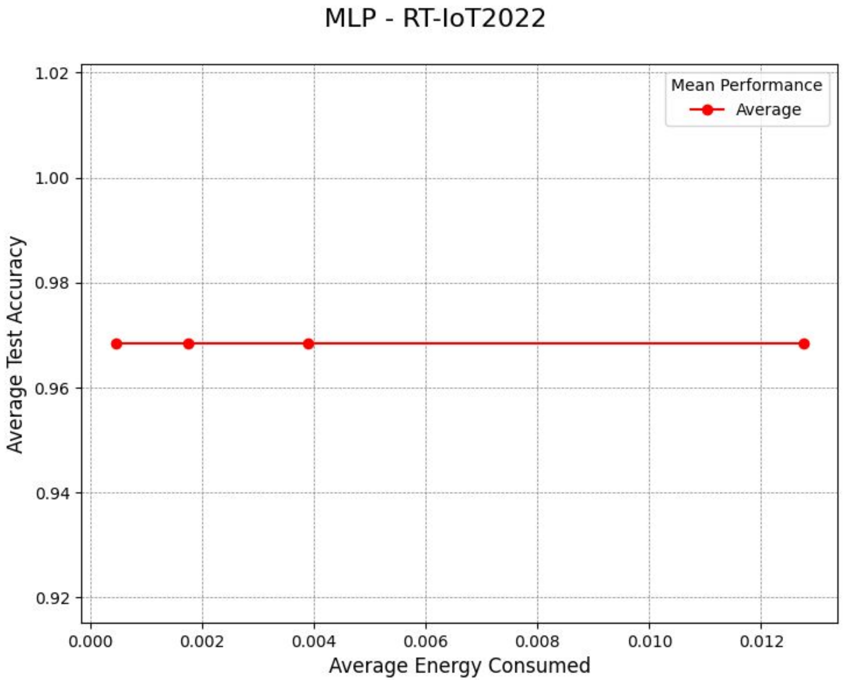
The Results - KNN



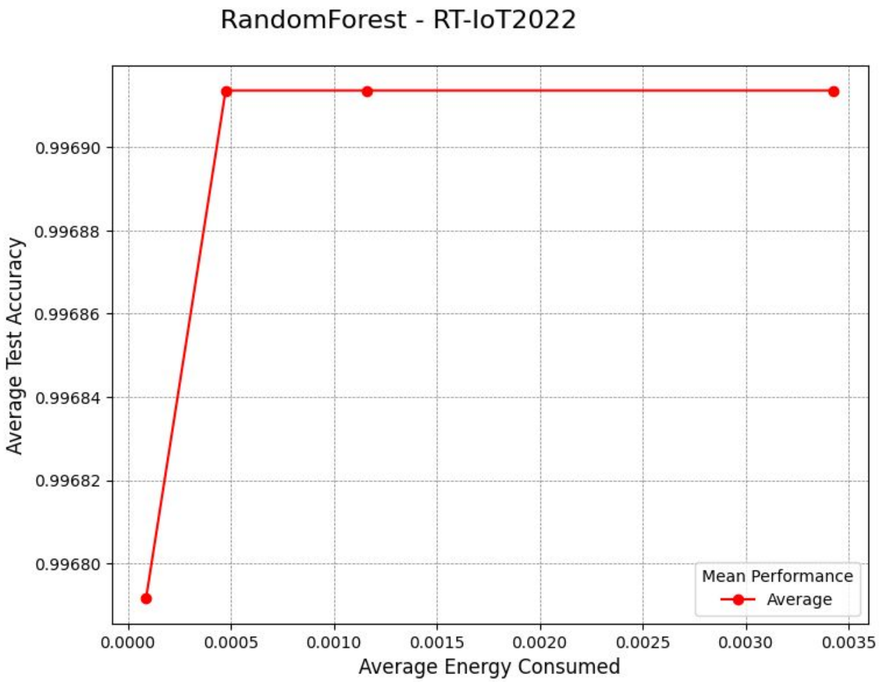


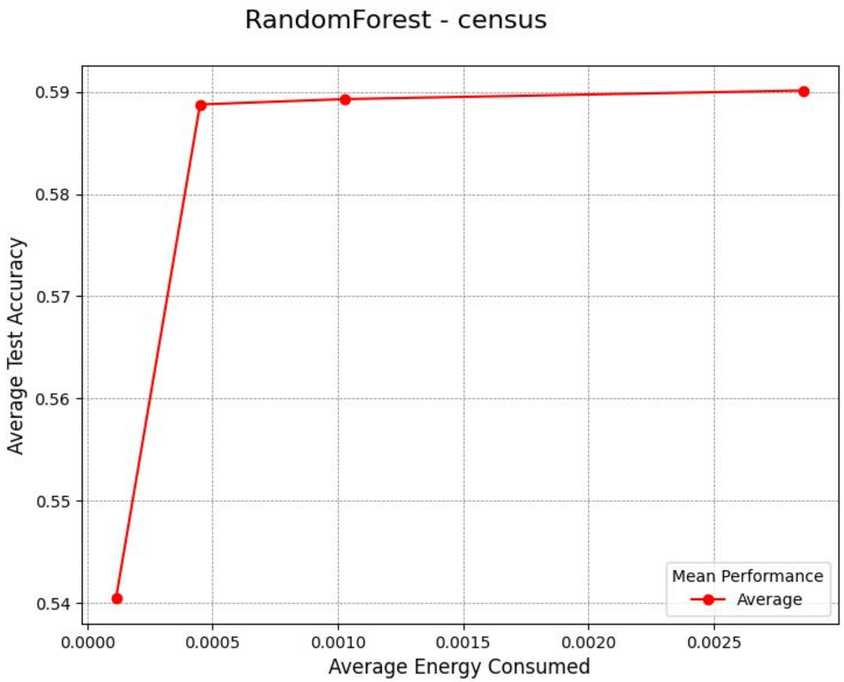
The Results - MLPs





The Results - Random Forest


RandomForest - RT-IoT2022


RandomForest - census

Findings

- Most algorithms showed **minimal increase** in accuracy
- Testing additional hyperparameters increased energy consumption dramatically.
- Average increase in energy per additional grid search is **309.64%**
- Average increase in accuracy per additional grid search is **0.01%**

Conclusion

- Can we say Grid Search is inefficient?
    - Mostly yes!

- There are clear points of diminishing/no returns

- Unless
    - any increase in accuracy vital to you, Or
    - you maybe don’t care about energy

Experiment Code

- Code is available at our repository: github.com/abdomash/ml-energy-consumption
- You can run the experiment on your device
- Setup guide in the README file