MachineLearnAthon - Microlecture Hyperparameter Tuning

Recorded by Lara Kuhlmann

MachineLearnAthon
A project Co-funded by the Erasmus+ programme of the European Union

Learning outcomes of today

After successfully completing this micro-lecture, you are able to….

- Understand the role and importance of hyperparameters in machine learning models
- Apply basic hyperparameter tuning techniques
- Understand the principles behind advanced methods

Agenda for today

- Hyperparameters
- Hyperparameter tuning
- Grid Search
- Random Search
- Hyperband optimization

Hyperparameters

- Hyperparameters are inputs to a machine learning algorithm that govern how the algorithm's perfomances generalizes to new, unseen data.[1]
- Are not learned from data, unlike model parameters
- Examples: learning rates, amount of regularization

[1] Li et al. 2018

Co-funded by the Erasmus+ programme of the European Union

Hyperparameter tuning

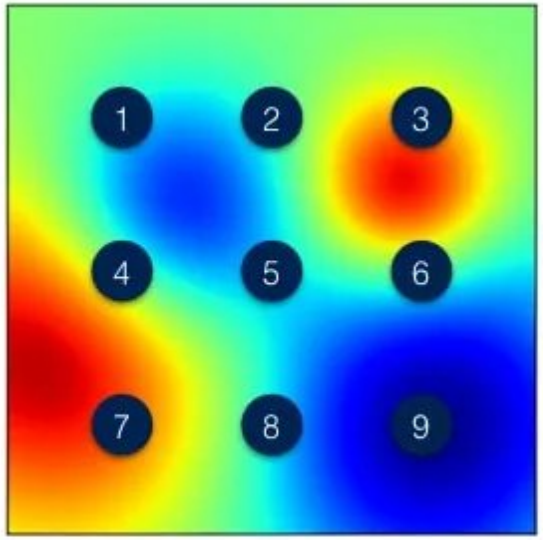
- Hyperparameter tuning is the process of determining optimal hyperparameters to measure performance of a machine learning model.[1]
- Goal: Find the highest possible model performance through a few sequential queries.[2]
- Challenge: Complexity increases with the number of hyperparameters and their interactions.[3]
- Tuning methods: Manual adjustment, Automated Algorithms

[1] Gerouani & Bouneffa. 2024
[2] Nguyen. 2019
[3] Sievert et al. 2019; Li et al. 2018

Co-funded by the Erasmus+ programme of the European Union

Grid Search

- Methods: Systematic search through all combinations in a given parameter range.[1]
- Advantages: Easy to implement and to understand.[2]
- Disadvantages: High computing effort, inefficient with large search spaces.[3]

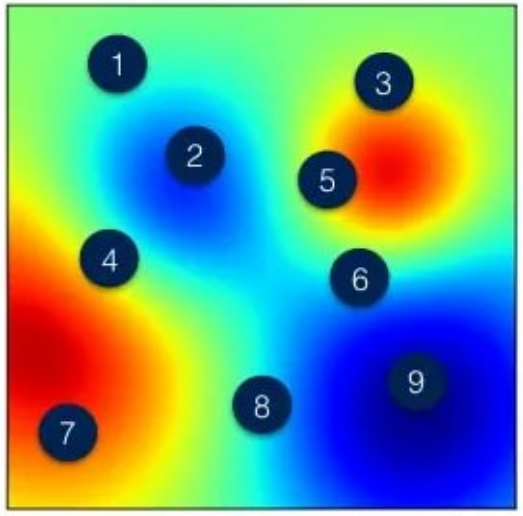


Grid Search

[1] Raschka et al. 2020
[2][3] Juilian & Devipriya. 2024
Iimg:https://pub.aimind.so/understanding-hyperparameter-optimization-techniques-4a39d0494612

Random Search

- Methods: Random selection and evaluation of configurations within the parameter range.[1]
- Advantages: Flexible and fault-tolerant, experiments can be stopped or supplemented at any time.[2]
- Disadvantages: No complete overview of the entire parameter range.[3]



Random Search

[1] Raschka et al. 2020
[2] Bergstra & Bengio. 2012
[3] Juilian & Devipriya. 2024
[img:https://pub.aimind.so/understanding-hyperparameter-optimization-techniques-4a39d0494612

technische universität dortmund
TU Delft
Co-funded by the Erasmus+ programme of the European Union
Machine LearnAthon

Hyperband Optimization

- Principle:
  - Exploration vs. exploitation: Hyperband tests many different configurations (exploration), with the most promising ones being pursued further (exploitation).[1]
  - Successive Halving: Several configurations receive certain resources at the beginning, weak configurations are gradually sorted out, better configurations receive more resources.[2]
- Functionality:
  - Generation of multiple, random configurations of the hyperparameters.[1]
  - Distributes a fixed amount of resources to the configurations.[2]
  - Each configuration is trained and evaluated for a certain number of iterations.[3]
  - Weak configurations are aborted at an early stage to save resources.[3]
  - Successful configurations are further optimized.[3]
  - Finally, the power of the individual brackets is averaged to give a total power.[2]

[1] Raschka et al. 2020
[2] Li et al. 2018
[3] Sievert et al. 2019
img:https://pub.aimind.so/understanding-hyperparameter-optimization-techniques-4a39d0494612

technische universität dortmund
TU Delft
Co-funded by the Erasmus+ programme of the European Union
Machine LearnAthon

Recap this lecture

After successfully completing this lecture, you are able to….

- Understand the role and importance of hyperparameters in machine learning models
- Apply basic hyperparameter tuning techniques
- Understand the principles behind advanced methods

The European Commission support for the production of this publication does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

This material was created with the assistance of an AI language model.

technische universität dortmund
TU Delft
Co-funded by the Erasmus+ programme of the European Union
Machine LearnAthon