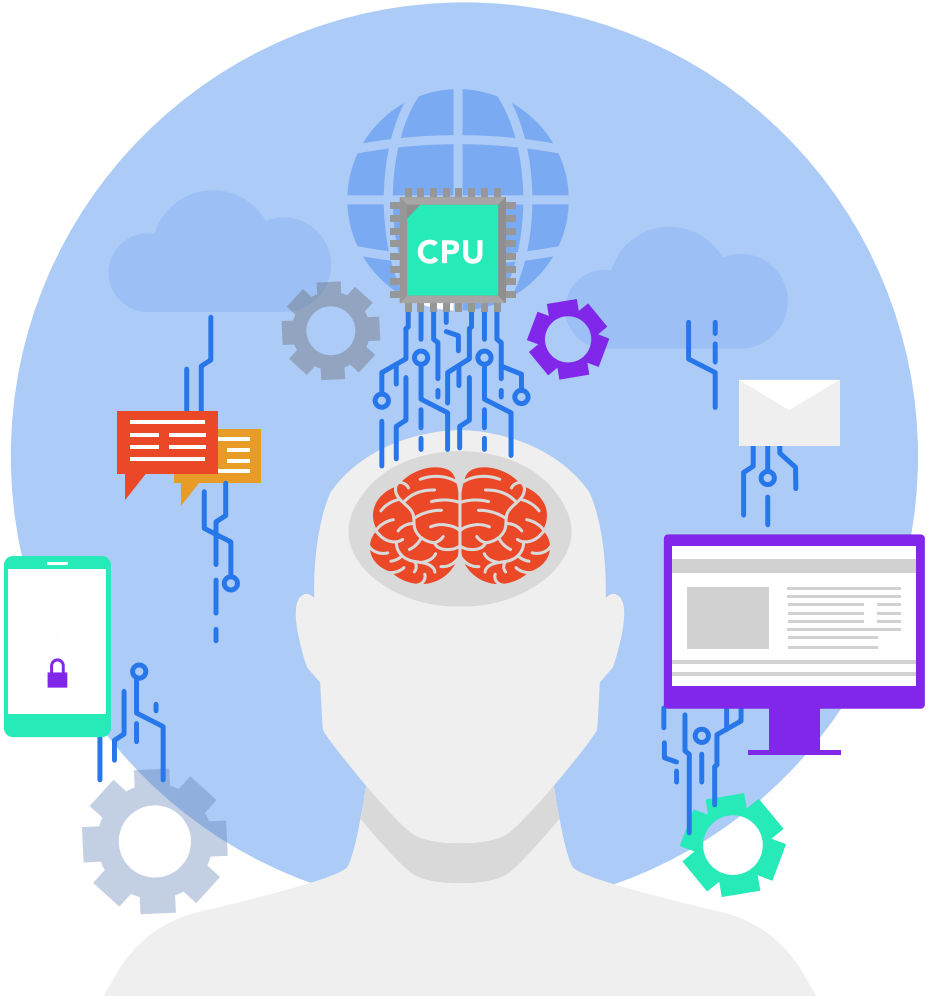


Hyper-parameter tuning

Agenda

01 **What is hyperparameter tuning?**

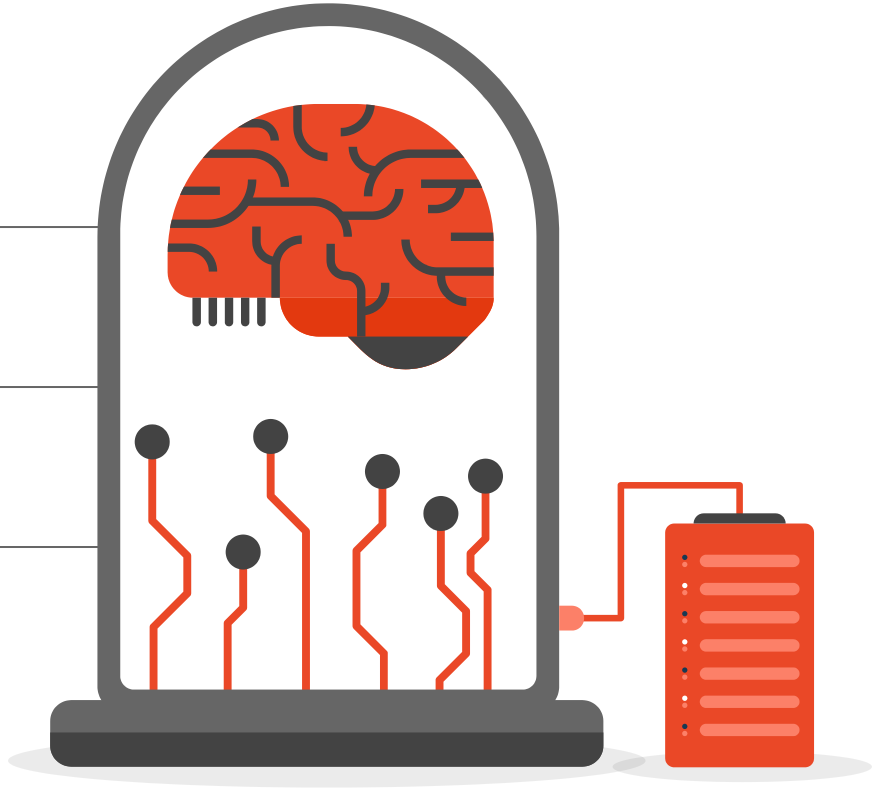
Why do we need it?

02 **Common Hyperparameters**

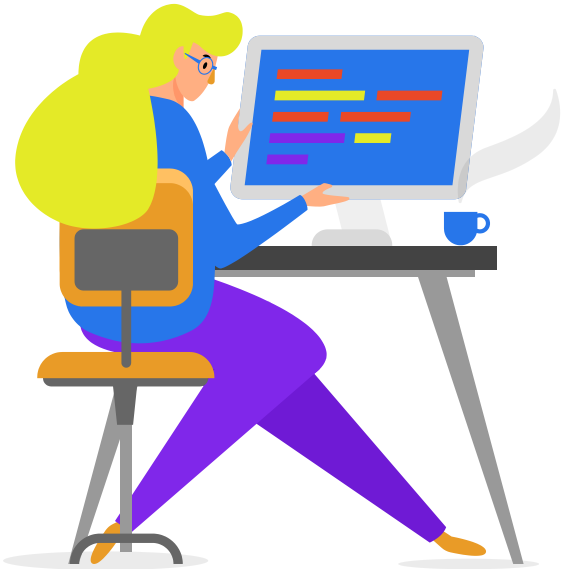
What hyperparameters are generally tuned?

03 **Hyperparameter optimization algorithms**

What algorithms are used to help us
In tuning hyperparameters?



What is hyperparameter tuning?



Hyperparameter Tuning

- Optimizing hyperparameters to minimize loss
- Generalized solution
- Reasonable timeframe
- No “one correct way”

What is hyperparameter tuning?

Hyperparameter Tuning

- Optimizing hyperparameters to minimize loss
- Generalized solution
- Reasonable timeframe
- No “one correct way”

Dataset

Model type

Comp. Power

Common hyperparameters

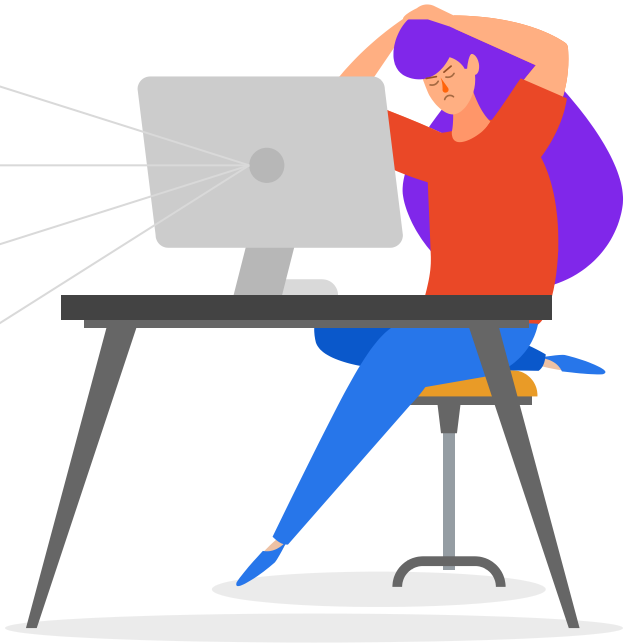
Learning rate
The rate of adjustment each step

Momentum
“Resistance” to rapid changes in values

Number of hidden layers
More isn’t always better

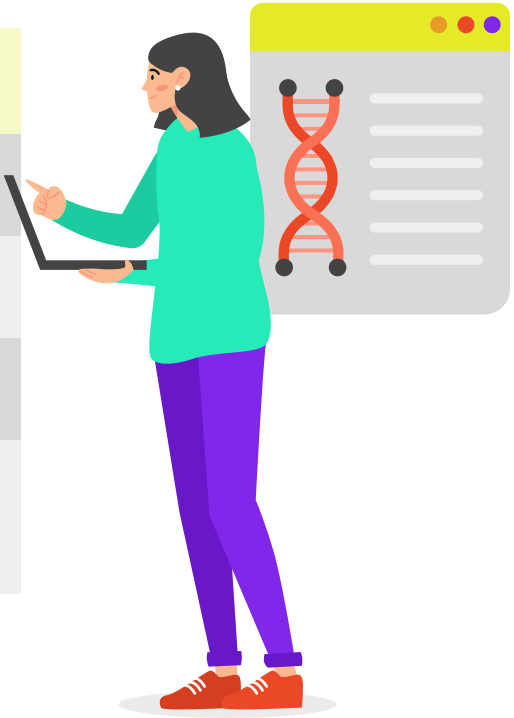
Number of neurons in each layer
Memorization vs Generalization

01
02
03
04

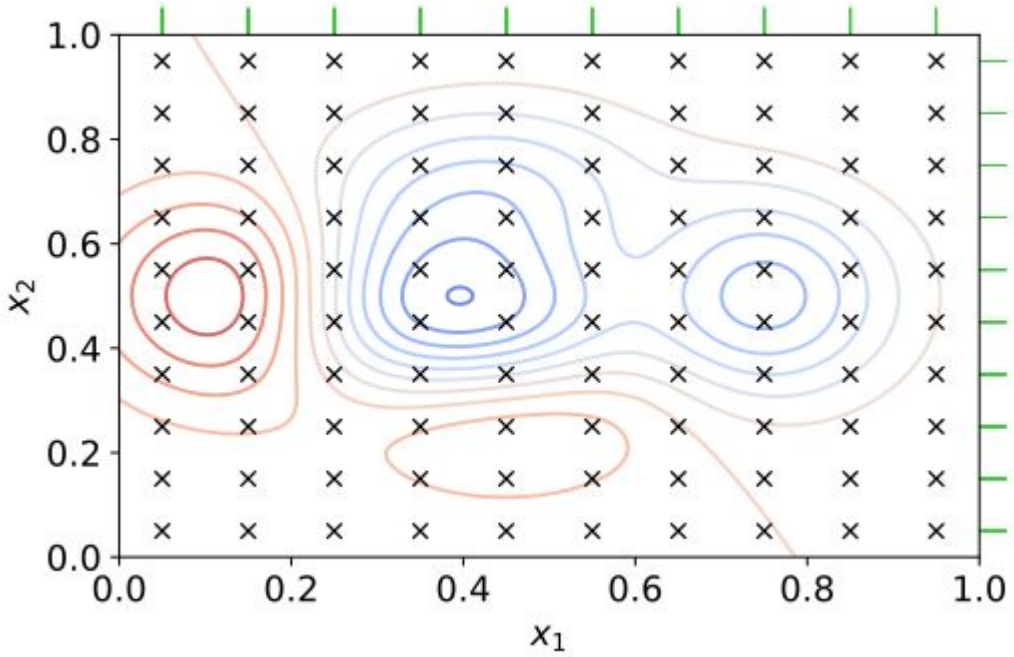


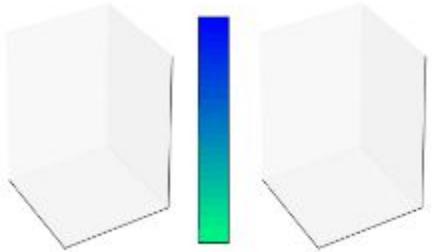
Hyperparameter optimization algorithms

Algorithms	
Manual	Manually setting parameters
Grid Search	Computes all possible combination
Random Search	Randomly picks N combinations
Automated Hyperparameter Tuning	E.g. Bayesian Optimization and Genetic Algorithms

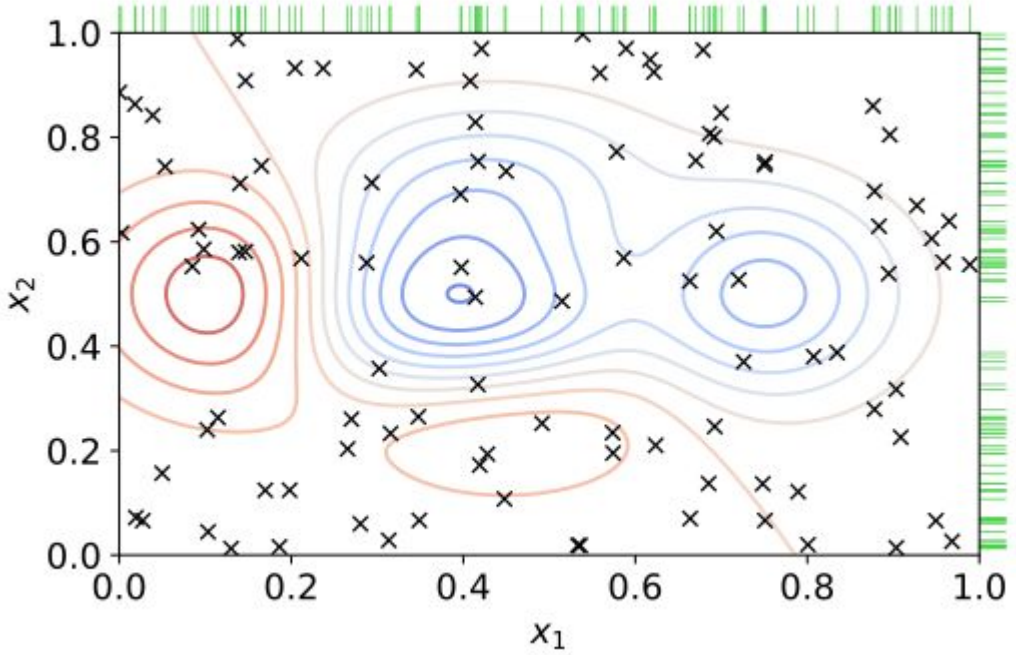


Grid search

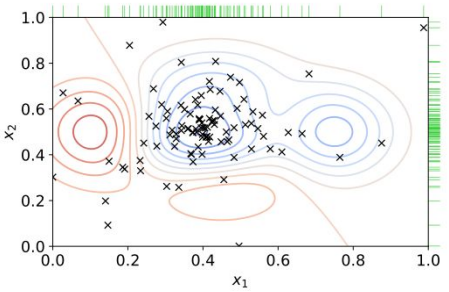




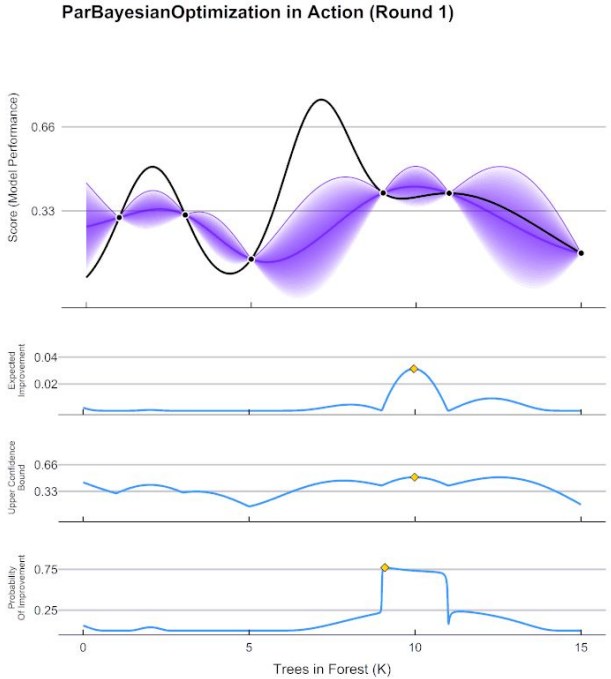
Random search



Automated Hyperparameter Tuning (Bayesian Optimization)

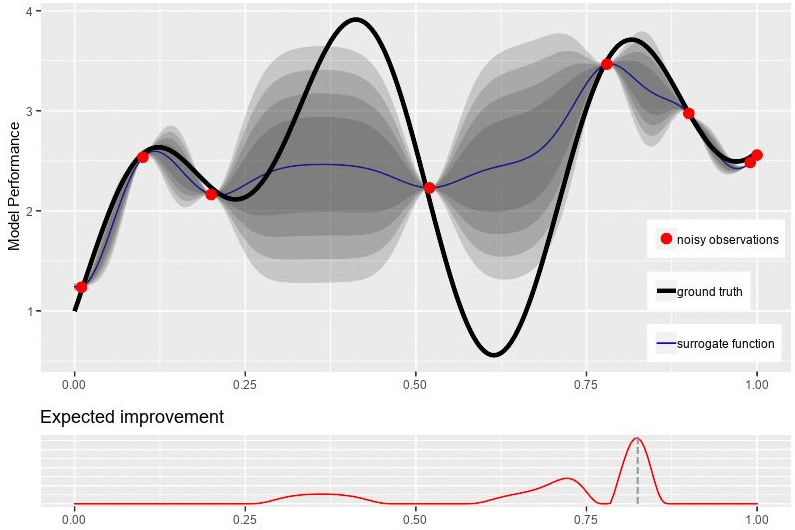


- Two key differences:
  - Uses a probabilistic surrogate function
  - Uses previous results
- Has two extra parts:
  - Surrogate function (SF)
  - Acquisition function (AF)
- Can use most probability functions
  - Surrogate: Gaussian probability
  - Acquisition:
    - Probability of improvement
    - GP upper confidence bound
    - Predictive entropy search
    - A portfolio of multiple acquisition strategies



Automated Hyperparameter Tuning (Bayesian Optimization) cont.

- Consists of 5 steps:
  1. Initialize the surrogate function (SF)
  2. Use the Acquisition function (AF) to find the best performing set, X
  3. Update the acquisition function based on SF
  4. Calculate OF(X)
  5. Calculate the difference between OF(X) and SF(X) and update SF based on result



Sources:

- T. Yu and H. Zhu, “Hyper-parameter optimization: A review of algorithms and applications,” *arXiv.org*, 12-Mar-2020. [Online]. Available: https://arxiv.org/abs/2003.05689. [Accessed: 12-Oct-2022].
- “What is hyperparameter tuning?,” *Anyscale*. [Online]. Available: https://www.anyscale.com/blog/what-is-hyperparameter-tuning. [Accessed: 12-Oct-2022].
- P. P. Ippolito, “Hyperparameters optimization,” *Medium*, 26-Sep-2019. [Online]. Available: https://towardsdatascience.com/hyperparameters-optimization-526348bb8e2d. [Accessed: 12-Oct-2022].
- T. Nunes, “Hyperparameter tuning,” *Medium*, 06-Aug-2020. [Online]. Available: https://medium.com/analytics-vidhya/hyperparameter-tuning-8ca311b16057. [Accessed: 12-Oct-2022].