

Artificial Intelligence (CS303)

Lecture 9: Automated Machine Learning

Hints for this lecture

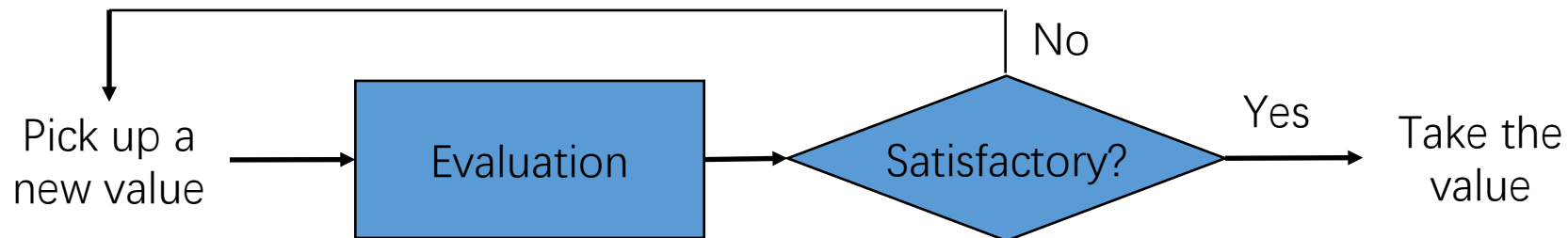
- Most (if not all) machine machine learning tasks involves hyper-parameters, which need to be tuned.

Hyper-parameters of ML Tasks

- Supervised Learning
 - SVM: Kernel Parameters and Regularization Parameters
 - Neural Networks: Number of hidden nodes, activation functions, network architecture
 - Decision Tree: Branching factor, Height of the tree
- Unsupervised Learning
 - Clustering: Number of clusters

Tuning Hyper-parameters

- How to tune the hyper-parameters?
 - No analytical solution to the optimization problem, since evaluating a value of hyper-parameter involves **estimating the generalization performance** of the resultant model.
 - Trial-and-error is the only methodology.



Tuning Hyper-parameters

- Grid Search
 - Too costly
- More efficient ways?
 - Use heuristic Search (e.g., using Black-Box optimization algorithms)
 - Sometimes, good surrogate of generalization is available to accelerate the evaluation
 - Take kernelized version of linear regression as an example
 - Generalization could be estimated with leave-one-out cross-validation.
 - With $\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & A_{22} \end{bmatrix}^{-1}$, we can exactly calculate the LOOCV error with only one matrix inversion.

The End of Machine Learning Section