

Chapter1

1. How would you define Machine Learning?

A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E.

2. Can you name four types of problems where it shines?

Problems for which existing solutions require a lot of hand-tuning or long lists of rules: one Machine Learning algorithm can often simplify code and perform better.

Complex problems for which there is no good solution at all using a traditional approach: the best Machine Learning techniques can find a solution.

Fluctuating environments: a Machine Learning system can adapt to new data.

Getting insights about complex problems and large amounts of data.

3. What is a labeled training set?

A training set where the training data you feed to the algorithm includes the desired solutions.

4. What are the two most common supervised tasks?

Classification and Regression.

5. Can you name four common unsupervised tasks?

Clustering, Anomaly detection and novelty detection, Visualization and dimensionality reduction , Association rule learning.

6. What type of Machine Learning algorithm would you use to allow a robot to walk in various unknown terrains?

Reinforcement Learning.

7. What type of algorithm would you use to segment your customers into multiple groups?

Unsupervised learning.

8. Would you frame the problem of spam detection as a supervised learning problem or an unsupervised learning problem?

Supervised learning.

9. What is an online learning system?

A learning system where you train the system incrementally by feeding it data instances sequentially, either individually or by small groups called mini-batches.

10. What is out-of-core learning?

Training systems on huge datasets that cannot fit in one machine's main memory.

11. What type of learning algorithm relies on a similarity measure to make predictions?

Instance-based learning.

12. What is the difference between a model parameter and a learning algorithm's hyperparameter?

Hyperparamete is a parameter of a learning algorithm and not affected by the learning algorithm itself. It must be set prior to the training and remains constant during training.

A model prarmeter is for a model.

13. What do model-based learning algorithms search for? What is the most common strategy they use to succeed? How do they make predictions?

They search for the model parameter values that minimize a cost function.

By minimizing the cost function or something like that.

Using the value of new instance and applying the model.

14. Can you name four of the main challenges in Machine Learning?

Insufficient Quantity of Training Data.

Nonrepresentative Training Data.

Poor-Quality Data.

Irrelevant Features.

15. If your model performs great on the training data but generalizes poorly to new instances, what is happening? Can you name three possible solutions?

Overfitting the training data.

To simplify the model by selecting one with fewer parameters.

To gather more training data.

To reduce the noise in the training data.

16. What is a test set and why would you want to use it?

A test set is a dataset to evaluate your model.

For evaluating the performance of your model.

17. What is the purpose of a validation set?

To select the best model.

18. What can go wrong if you tune hyperparameters using the test set?

The model may work badly when put into production.

19. What is repeated cross-validation and why would you prefer it to using a single validation set?

Each model is evaluated once per validation set, after it is trained on the rest of the data.

By averaging out all the evaluations of a model, we get a much more accurate measure of its performance.