Exam practice questions

barbaras

May 2024

1 PCA and Visualization

- What is the main goal of PCA? Is PCA an instance of supervised or unsupervised learning? Is PCA different from clustering? If so, how?

2 Recommender Systems

- What are the main differences between collaborative and content based filtering?

3 Clustering

- Describe the k-means algorithm. How is it different from clustering with GMM?

4 Gaussian Mixture Models

- The GMM can be applied in both supervised, semi-supervised, and unsupervised settings. The GMM has an analytical solution in the case of supervised learning, but not for semi- and unsupervised settings. However, in such cases we can still have a numerical solution. Which algorithm is used in these cases? How does it work? Is it guaranteed to find an optimal solution?

5 Practical Considerations on Supervised ML

- What is the generalization gap and what strategies can be used to reduce it?

6 Trees

- Describe the role of entropy and/or Gini impurity in decision trees.

7 Bagging

- Describe how bagging (bootstrap aggregating) works.
- What usually happens to the training and test error when we use bagging?

8 Boosting

- Describe how boosting works.

9 Support Vector Machines (SVMs)

- What is a maximal margin classifier, and how does it work?

10 Reinforcement Learning

- Describe the role of rewards in reinforcement learning.
- What is a Bellman equation (or dynamic programming) in reinforcement learning?