

Questions Lecture 2 Paper 4 Paper 5

4712099

Lecture 2

1. Minimizing the KL divergence corresponds to:
 - a. Minimizing the cross entropy between distributions
 - b. maximize maximum likelihood
 - c. Minimizing of the negative log-likelihood function
 - d. A or B

Answer A. Minimizing the cross entropy between distributions

Paper 4 : CNN Features off-the-shelf an Astounding Baseline for Recognition

1. The CNN features are compared against state of the art retrieval pipelines for visual instance retrieval. Which of the following methods use dictionaries that are trained on similar train and test dataset?
 - a. VLAD, BoW, IFV, Hamming Embedding, BoB
 - b. BoW+CNN only
 - c. CNN representation +SVM
 - d. Only C

Answer A. VLAD, BoW, IFV, Hamming Embedding, BoB

Paper 5: Large-Scale Machine Learning with Stochastic Gradient Descent

1. The excess error is decomposed into three terms - The approximation error, the estimation error, the optimization error.
Which is true about the three terms mentioned?
 - a. The approximation error is reduced by choosing a small family of functions
 - b. The estimation error can be reduced by choosing a smaller family of functions and by increasing the size of the training data set simultaneously
 - c. Optimization error can be reduced by running the optimizer longer
 - d. Both a and c.

Answer: C. Optimization error can be reduced by running the optimizer longer.