

There will be 3 types of questions roughly described as: theory questions, small problems, and large problems. Here are some examples:

Theory questions:

1. What is the difference between supervised and unsupervised learning
2. Explain one issue arising when applying gradient descent algorithms on non-convex problems
3. What is the difference between batch and stochastic gradient algorithm
4. What is curse of dimensionality
5. What is a rank of a matrix

Smaller problems:

6. You are given a neural network with 10 inputs, 3 hidden neurons, and one output neuron. How many weights does this neural network have?
7. Given a scatterplot with 10 positive and 10 negative examples, decide the class label of an unlabeled example in the scatterplot (the scatterplot would be provided) for k-nearest neighbor for  $k = 1$  and  $k = 3$ .
8. Two of the features in the data set used for regression are identical. What is a potential issue with using the closed-form solution for linear regression on such data? Propose one way to resolve it.
9. Training a predictor on  $n$  training examples takes  $n^2$  seconds and testing on  $n$  test examples takes  $n$  seconds. What is the total time required to perform the 5 cross validation on a data set with 100 labeled examples?

Larger problems

10. Given a loss function  $\text{loss} = \max(0, y * w' * x)$ , and assuming  $w' = [1 \ -2]$ , what is an updated weight obtained by the stochastic gradient descent upon receiving a new example  $x' = [0 \ 1]$ ,  $y = +1$ . Assume the learning rate is 0.01.
11. What is the time complexity (using big-Oh notation) of: a) adding two vectors with  $n$  elements, b) multiplying 2 square matrices of size  $n \times n$ , c) finding a nearest neighbor given a training data of size  $n \times m$ , d) doing one update of batch-mode gradient descent of a neural network with one hidden layer on training data of size  $n \times m$ .