

1
point

1.

Suppose you are training a logistic regression classifier using stochastic gradient descent. You find that the cost (say, $\text{cost}(\theta, (x^{(i)}, y^{(i)}))$, averaged over the last 500 examples), plotted as a function of the number of iterations, is slowly increasing over time. Which of the following changes are likely to help?

- ☐ This is not possible with stochastic gradient descent, as it is guaranteed to converge to the optimal parameters θ .
- ☒ Try halving (decreasing) the learning rate α , and see if that causes the cost to now consistently go down; and if not, keep halving it until it does.
- ☐ Use fewer examples from your training set.
- ☐ Try averaging the cost over a smaller number of examples (say 250 examples instead of 500) in the plot.

1
point

2.

Which of the following statements about stochastic gradient descent are true? Check all that apply.

- ☒ You can use the method of numerical gradient checking to verify that your stochastic gradient descent implementation is bug-free. (One step of stochastic gradient descent computes the partial derivative $\frac{\partial}{\partial \theta_j} \text{cost}(\theta, (x^{(i)}, y^{(i)}))$.)
- ☒ Before running stochastic gradient descent, you should randomly shuffle (reorder) the training set.

- ☐ Suppose you are using stochastic gradient descent to train a linear regression classifier. The cost function $J(\theta) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$ is guaranteed to decrease after every iteration of the stochastic gradient descent algorithm.

- ☐ In order to make sure stochastic gradient descent is converging, we typically compute $J_{\text{train}}(\theta)$ after each iteration (and plot it) in order to make sure that the cost function is generally decreasing.

1
point

3.

Which of the following statements about online learning are true? Check all that apply.

Large Scale Machine Learning

Quiz, 5 questions

- ☐ One of the advantages of online learning is that if the function we're modeling changes over time (such as if we are modeling the probability of clicking on different URLs, and user tastes/preferences are changing over time), the online learning algorithm will automatically adapt to these changes.

- ☐ When using online learning, you must save every new training example you get, as you will need to reuse past examples to re-train the model even after you get new training examples in the future.

- ☐ Online learning algorithms are usually best suited to problems where we have a continuous/non-stop stream of data that we want to learn from.

- ☐ Online learning algorithms are most appropriate when we have a fixed training set of size m that we want to train on.

1
point

4.

Assuming that you have a very large training set, which of the following algorithms do you think can be parallelized using map-reduce and splitting the training set across different machines? Check all that apply.

- ☐ Logistic regression trained using stochastic gradient descent.
- ☐ Computing the average of all the features in your training set $\mu = \frac{1}{m} \sum_{i=1}^m x^{(i)}$ (say in order to perform mean normalization).
- ☐ Logistic regression trained using batch gradient descent.
- ☐ Linear regression trained using stochastic gradient descent.

Large Scale Machine Learning

Quiz, 5 questions

5.

Which of the following statements about map-reduce are true? Check all that apply.

- ☐ If you have only 1 computer with 1 computing core, then map-reduce is unlikely to help.
- ☐ When using map-reduce with gradient descent, we usually use a single machine that accumulates the gradients from each of the map-reduce machines, in order to compute the parameter update for that iteration.
- ☐ Because of network latency and other overhead associated with map-reduce, if we run map-reduce using N computers, we might get less than an N -fold speedup compared to using 1 computer.
- ☐ If we run map-reduce using N computers, then we will always get at least an N -fold speedup compared to using 1 computer.

☐ I, **Jun-Chieh Wang**, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account. Learn more about Coursera's Honor Code

Submit Quiz



Large Scale Machine Learning

Quiz, 5 questions