# Large Scale Machine Learning

Cuestionario, 5 preguntas

1 point

1.

Suppose you are training a logistic regression classifier using stochastic gradient descent. You find that the cost (say, $cost(\theta, (x^{(i)}, y^{(i)}))$, 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 $\theta$.
- [x] Try halving (decreasing) the learning rate $\alpha$, and see if that causes the cost to now consistently go down; and if not, keep halving it until it does.
- [ ] Try averaging the cost over a smaller number of examples (say 250 examples instead of 500) in the plot.
- [ ] Use fewer examples from your training set.

---

1 point

2.

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

- [x] 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.
- [x] 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 $\frac{\partial}{\partial\theta_j}cost(\theta, (x^{(i)}, y^{(i)}))$.)
- [ ] 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.

- [ ] Online learning algorithms are most appropriate when we have a fixed training set of size $m$ that we want to train on.
- [x] 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 users 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.
- [x] Online learning algorithms are usually best suited to problems were we have a continuous/non-stop stream of data that we want to learn from.

---

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.

- [ ] Linear regression trained using stochastic gradient descent.
- [ ] Logistic regression trained using stochastic gradient descent.
- [x] 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).
- [x] Logistic regression trained using batch gradient descent.

---

1 point

5.

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

- [x] 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.
- [x] If you have only 1 computer with 1 computing core, then map-reduce is unlikely to help.
- [ ] Linear regression and logistic regression can be parallelized using map-reduce, but not neural network training.
- [x] 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.

---

Entiendo que enviar trabajo que no es mío resultará en la desaprobación permanente de este curso y la desactivación de mi cuenta de Coursera.

Obtén más información sobre el Código de Honor de Coursera

Ingresa tu nombre legal

Submit Quiz