WEEK 3 Quiz

Total points 13/19

Cost Function, Gradient Descent, Linear Regression, Logistic Regression, Over-fitting, Under-fitting, Bias Variance Trade-off

Total Quiz is of 20 marks

13 of 19 points

✓ Learning rate is an hyper parameter in Gradient Descent used to update the parameters in the optimization problem. * 1/1

- ○ False
- ◉ True ✓

✗ Gradient of a continuous and differentiable function is * 0/1

- [x] is zero at a minimum ✓
- [ ] is non-zero at a maximum
- [ ] is zero at a saddle point
- [x] decreases as you get closer to the minimum ✓

Let f be some function so that $f(\theta_0,\theta_1)$ outputs a number. For this problem, f is some arbitrary/unknown smooth function (not necessarily the cost function of linear regression, so f may have local optima).Suppose we use gradient descent to try to minimize $f(\theta_0,\theta_1)$ as a function of θ_0 and θ_1. Which of the following statements are true? (Check all that apply.) * 1/1

- [ ] Even if the learning rate α is very large, every iteration of gradient descent will decrease the value of $f(\theta_0,\theta_1)$.
- [x] If the learning rate is too small, then gradient descent may take a very long time to converge. ✓
- [x] If θ_0 and θ_1 are initialized at a local minimum, then one iteration will not change their values. ✓
- [ ] If θ_0 and θ_1 are initialized so that $\theta_0 = \theta_1$, then by symmetry (because we do simultaneous updates to the two parameters), after one iteration of gradient descent, we will still have $\theta_0=\theta_1$.

You observe the following while fitting a linear regression to the data: As you increase the amount of training data, the test error decreases and the training error increases. The train error is quite low (almost what you expect it to), while the test error is much higher than the train error. What do you think is the main reason behind this behavior. Choose the most probable option. * 0/1

- ( ) High variance
- (•) High model bias ✗
- ( ) High estimation bias
- ( ) None of the above

✓ Which of the following sentence is FALSE regarding regression? * 1/1

- ○ It relates inputs to outputs.
- ○ It is used for prediction.
- ○ It may be used for interpretation.
- ◉ It discovers causal relationships ✓

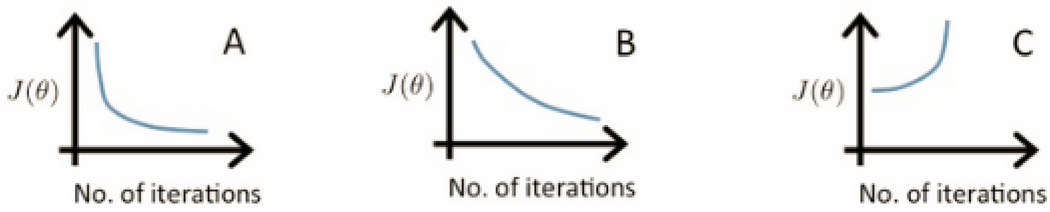
✗ Overfitting is more likely when you have huge amount of data to train? * 0/1

- ◉ True ✗
- ○ False

✓ Which of the following statement is true about outliers in Linear regression? * 1/1

- ◉ Linear regression is sensitive to outliers ✓
- ○ Linear regression is not sensitive to outliers
- ○ Can’t say
- ○ None of these

Cost function vs Number of Iterations



✓ Suppose l1, l2 and l3 are the three learning rates for A,B,C respectively. Which of the following is true about l1, l2 and l3? (This question is related to above image) * 1/1

- (●) l2 < l1 < l3 ✓
- ( ) l1 > l2 > l3
- ( ) l1 = l2 = l3
- ( ) None of these

✓ What do you expect will happen with bias and variance as you increase the size of training data? * 1/1

- ( ) Bias increases and Variance increases
- ( ) Bias decreases and Variance increases
- ( ) Bias decreases and Variance decreases
- (●) Bias increases and Variance decreases ✓

Suppose, you got a situation where you find that your linear regression model is under fitting the data. In such situation which of the following options would you consider? 1. I will add more variables 2. I will start introducing polynomial degree variables 3. I will remove some variables * 0/1

- 1 and 2
- 2 and 3
- 1 and 3
- 1, 2 and 3

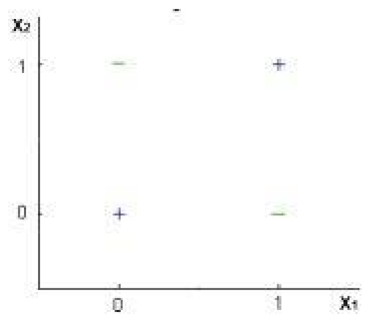
Is it possible to apply a logistic regression algorithm on a 3-class Classification problem? * 1/1

- True
- False

Choose which of the following options is true regarding One-Vs-All method in Logistic Regression. * 1/1

- We need to fit n models in n-class classification problem
- We need to fit n-1 models to classify into n classes
- We need to fit only 1 model to classify into n classes
- None of these

Data for the next question



Can a Logistic Regression classifier do a perfect classification on the below data? Note: You can use only X1 and X2 variables where X1 and X2 can take only two binary values(0,1). * 1/1

- True
- False ✓
- Can't say
- None of these

Regarding bias and variance, which of the following statements are true? (Here 'high' and 'low' are relative to the ideal model.) (i). Models which overfit are more likely to have high bias (ii). Models which overfit are more likely to have low bias (iii). Models which overfit are more likely to have high variance (iv). Models which overfit are more likely to have low variance * 0/1

- (i) and (ii)
- (ii) and (iii)
- (iii) and (iv)
- None of these ✗

The lines y = 1 and y = -1 are asymptotes to the Sigmoid Function * 1/1

- ( ) Yes,Both are asymptotes
- (•) Only y = 1 is asymptote ✓
- ( ) Only y = -1 is asymptote
- ( ) No, None of them are asymptotes to sigmoid function

How many hyper-parameters does linear regression have? (Kindly enter a number only without any other characters) * 0/1

2 ✗

If there are n (>2) classes how many classifiers would you have to train in One vs all classification ? (Just write answer, introductory texts are not required) * 1/1

n ✓

Linear Regression is more susceptible to ? * 1/1

- (•) Underfitting ✓
- ( ) Overfitting
- ( ) Both
- ( ) None

If the derivative of a differential and continuous function is 0, it implies global minima or maxima * 1/1

- True
- False

BEFORE SUBMITTING

The next page will lead you to accept the honor code and submit the User ID and Password. Make sure you have these ready.

THE USER ID AND PASSWORD FOR ALL QUIZZES/ASSIGNMENTS IS SAME THROUGHOUT COURSE

User Details 0 of 0 points

Enter your details carefully

Name *

Dhrubanka Dutta

User ID *

U011W5850DA

Password *

GN7Y34JB

Please type the following "I Accept the Honor Code and will not violate it in any possible way" *

I Accept the Honor Code and will not violate it in any possible way

This content is neither created nor endorsed by Google. - Terms of Service - Privacy Policy

Google Forms