

Classification and Representation

Logistic Regression Model

Multiclass Classification

Review

Solving the Problem of Overfitting

Video: The Problem of Overfitting  
9 min

Reading: The Problem of Overfitting  
3 min

Video: Cost Function  
10 min

Reading: Cost Function  
3 min

Video: Regularized Linear Regression  
10 min

Reading: Regularized Linear Regression  
3 min

Video: Regularized Logistic Regression  
8 min

Reading: Regularized Logistic Regression  
3 min

Review

Reading: Lecture Slides  
10 min

Quiz: Regularization  
5 questions

Programming Assignment: Logistic Regression  
3h

QUIZ • 10 MIN

Regularization

Submit your assignment

DUE Jul 26, 11:59 PM PDT    ATTEMPTS 3 every 8 hours

Receive grade

TO PASS 80% or higher

Regularization

TOTAL POINTS 5

1. You are training a classification model with logistic

1 point

regression. Which of the following statements are true? Check

all that apply.

- Start

☐ Adding a new feature to the model always results in equal or better performance on the training set.

☒ Introducing regularization to the model always results in equal or better performance on the training set.

☐ Adding many new features to the model helps prevent overfitting on the training set.

☐ Introducing regularization to the model always results in equal or better performance on examples not in the training set.
- 👍

👎

🗑

2. Suppose you ran logistic regression twice, once with  $\lambda = 0$ , and once with  $\lambda = 1$ . One of the times, you got

1 point

parameters  $\theta = \begin{bmatrix} 81.47 \\ 12.69 \end{bmatrix}$  and the other time you got

$\theta = \begin{bmatrix} 13.01 \\ 0.91 \end{bmatrix}$ . However, you forgot which value of

$\lambda$  corresponds to which value of  $\theta$ . Which one do you

think corresponds to  $\lambda = 1$ ?

- ☒  $\theta = \begin{bmatrix} 13.01 \\ 0.91 \end{bmatrix}$
- ☐  $\theta = \begin{bmatrix} 81.47 \\ 12.69 \end{bmatrix}$

3. Which of the following statements about regularization are

1 point

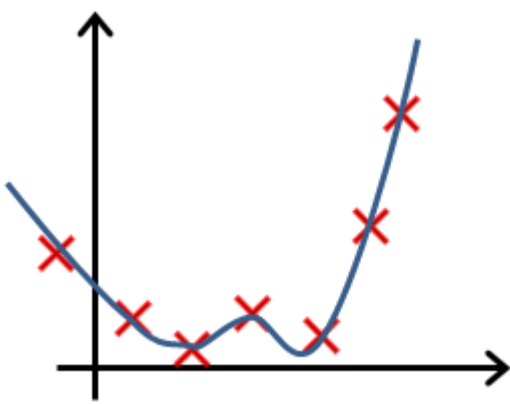
true? Check all that apply.

- ☒ Using too large a value of  $\lambda$  can cause your hypothesis to underfit the data.
- ☐ Because regularization causes  $J(\theta)$  to no longer be convex, gradient descent may not always converge to the global minimum (when  $\lambda > 0$ , and when using an appropriate learning rate  $\alpha$ ).
- ☐ Using a very large value of  $\lambda$  cannot hurt the performance of your hypothesis; the only reason we do not set  $\lambda$  to be too large is to avoid numerical problems.
- ☐ Because logistic regression outputs values  $0 \leq h_{\theta}(x) \leq 1$ , its range of output values can only be "shrunk" slightly by regularization anyway, so regularization is generally not helpful for it.

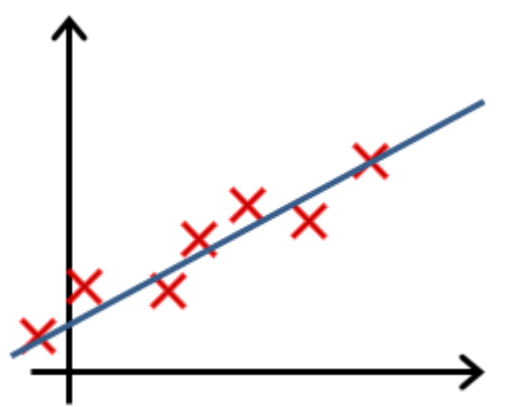
4. In which one of the following figures do you think the hypothesis has overfit the training set?

1 point

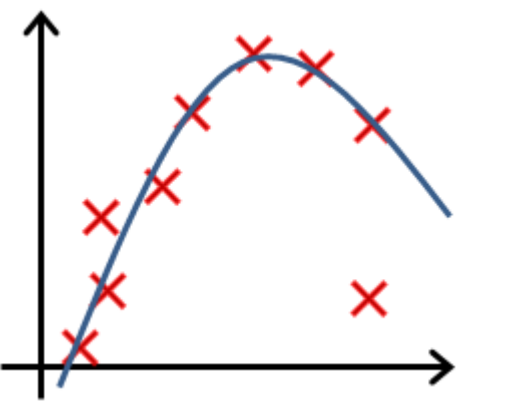
☒ Figure:



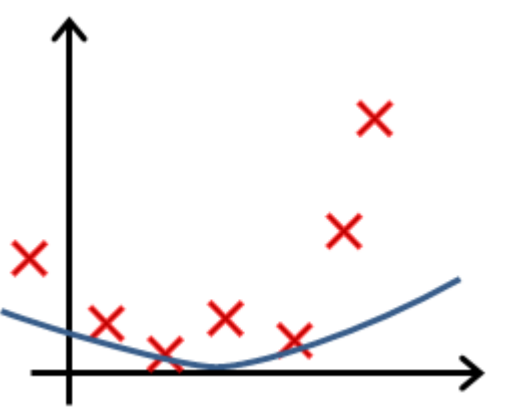
☐ Figure:



☐ Figure:



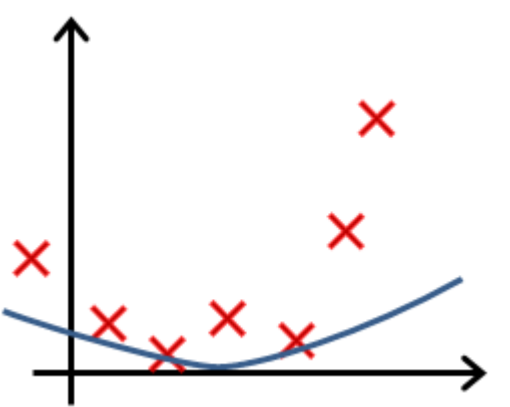
☐ Figure:



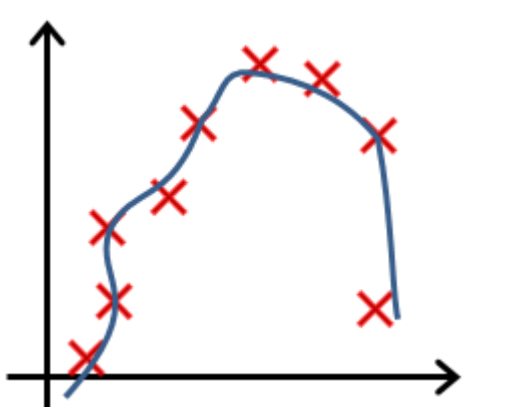
5. In which one of the following figures do you think the hypothesis has underfit the training set?

1 point

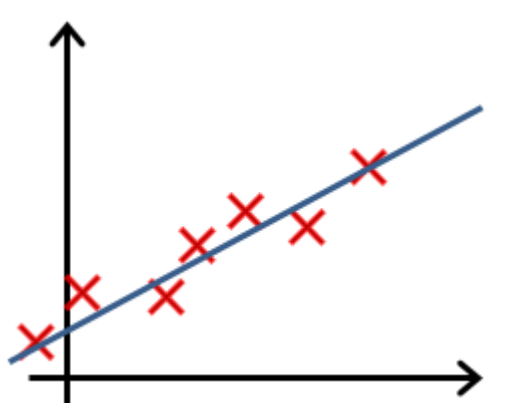
☒ Figure:



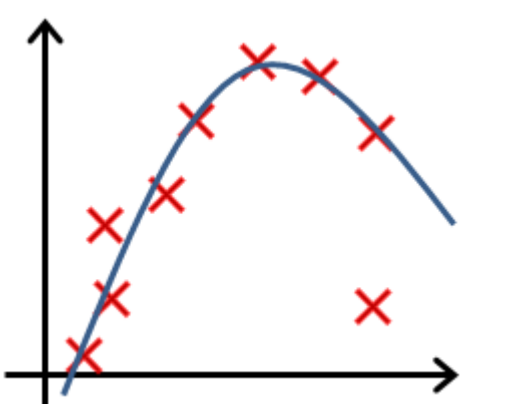
☐ Figure:



☐ Figure:



☐ Figure:



☒ I 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

👍

👎

🗑

Save

Submit