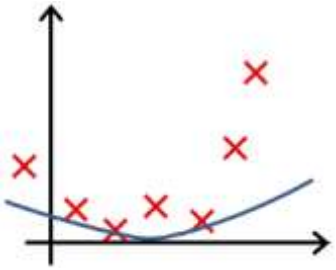
Quiz 7

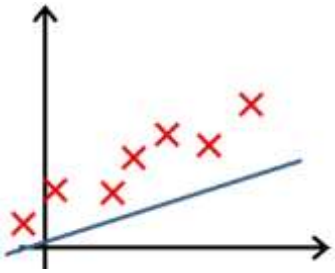
Question 1

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

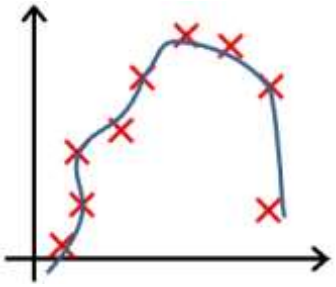
A ○



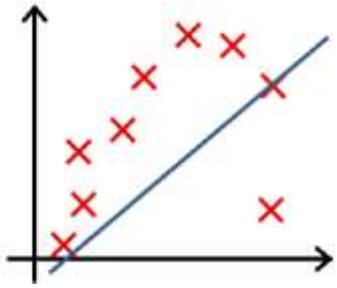
B ○



C ○



D ○

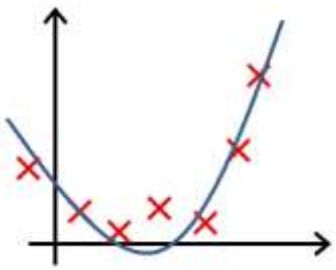


Quiz 7

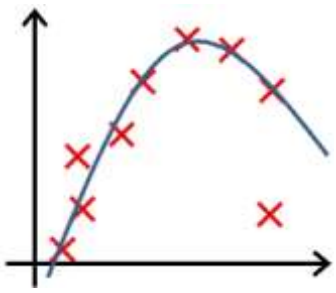
Question 2

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

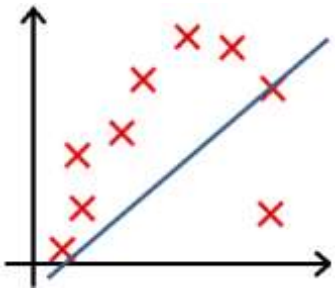
A ○



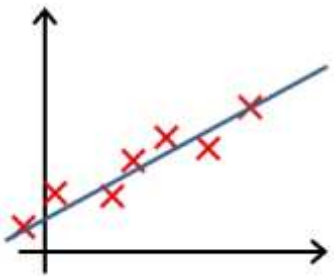
B ○



C ○



D ○



Question 3

You are training a classification model with logistic regression. Which of the following statements are true? Check all that apply.

Which of the following are true? Check all that apply.

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

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

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

D ☐ Adding many new features to the model makes it more likely to overfit the training set.

Quiz 7

Question 4

Suppose you ran logistic regression twice, once with $\lambda = 0$, and once with $\lambda = 1$. One of the times, you got parameters $\theta = \begin{bmatrix} 26.29 \\ 65.41 \end{bmatrix}$, and the other time you got $\theta = \begin{bmatrix} 2.75 \\ 1.32 \end{bmatrix}$. However, you forgot which value of λ corresponds to which value of θ. Which one do you think corresponds to $\lambda = 1$?

A ○ $\theta = \begin{bmatrix} 26.29 \\ 65.41 \end{bmatrix}$

B ○ $\theta = \begin{bmatrix} 2.75 \\ 1.32 \end{bmatrix}$

Question 5

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

A ☐ Using a very large value of λ cannot hurt the performance of your hypothesis; the only reason we do not set λ to be too large is to avoid numerical problems.

B ☐ Using too large a value of λ can cause your hypothesis to overfit the data; this can be avoided by reducing λ.

C ☐ 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 α).

D ☐ Using too large a value of λ can cause your hypothesis to underfit the data.