Week 3

Quiz 2-Regularization

Ques1. You are training a classification model with logistic regression. Which of the following statements are true? Check all that apply.

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

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

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

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

Answer:

A) Wrong. This will lead to overfitting
B) Correct.
C) Wrong. Can lead to underfitting
D) Wrong. Underfitting will lead to worse performance on egs not in the training set.

Ques2.

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

parameters $\theta = \begin{bmatrix} 74.81 \\ 45.05 \end{bmatrix}$, and the other time you got

$\theta = \begin{bmatrix} 1.37 \\ 0.51 \end{bmatrix}$. However, you forgot which value of

λ corresponds to which value of θ. Which one do you

think corresponds to $\lambda = 1$?

- (●) $\theta = \begin{bmatrix} 1.37 \\ 0.51 \end{bmatrix}$
- ( ) $\theta = \begin{bmatrix} 74.81 \\ 45.05 \end{bmatrix}$

Answer:
As when lambda = 1, we add the regularization term which will penalize when theta is big. Thus, when lambda = 1, theta will be relatively smaller than without regularization.

Ques3. Which of the following statements about regularization are true? Check all that apply.

A) Consider a classification problem. Adding regularization may cause your classifier to incorrectly classify some training examples (which it had correctly classified when not using regularization, i.e. when $\lambda=0$).

B) 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.

C) 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.

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

Answer:

A) Correct.
B) Wrong.
C) Wrong. Very large λ can lead to underfitting problem.
D) Wrong. Very large λ can lead to underfitting problem.

Ques4. **In** Which one of the following figures do you think the hypothesis has overfit the training set?

Figure:

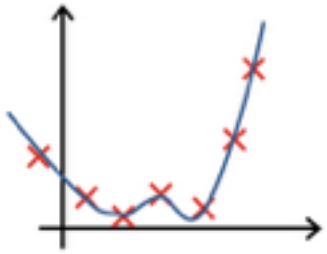


Figure:

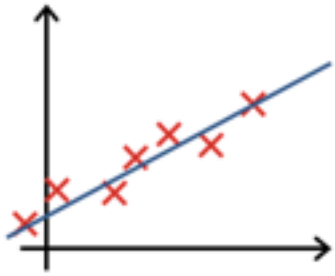


Figure:

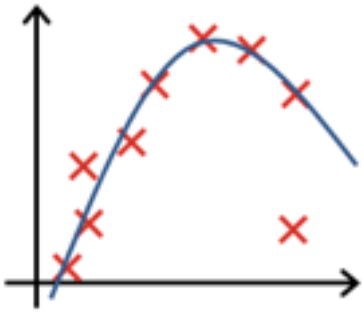
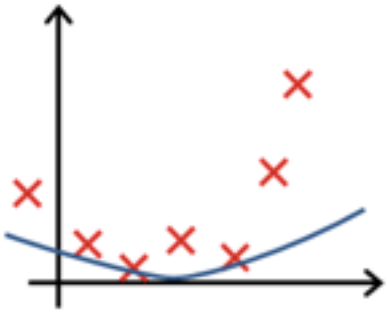


Figure:



Ques5. In Which one of the following figures do you think the hypothesis has underfit the training set?

Figure:

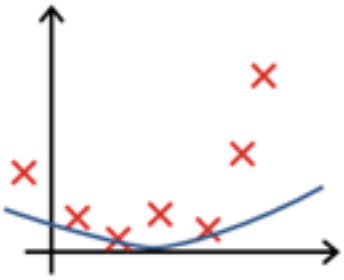


Figure:

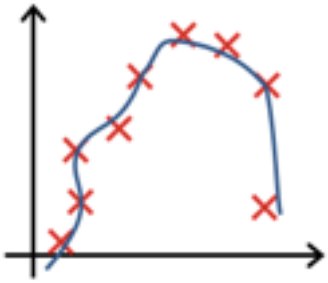


Figure:

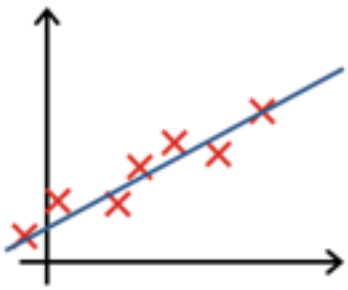


Figure: