Regularization

TOTAL POINTS 5

You are training a classification model with logistic regression. Which of the following statements are true? Check all that apply. 1 point

- [ ] Introducing regularization to the model always results in equal or better performance on examples not in 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 the training set.
- [x] Adding a new feature to the model always results in equal or better performance on the training set.

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

parameters $\theta = \begin{bmatrix} 74.81 \\ 45.05 \end{bmatrix}$, 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 $\lambda = 1$?

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

Which of the following statements about regularization are 1 point

true? Check all that apply.

- [ ] Using too large a value of λ can cause your hypothesis to overfit the data; this can be avoided by reducing λ.
- [ ] 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.
- [ ] 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.
- [x] 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 $\lambda = 0$).

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

λ corresponds to which value of θ. Which one do you think corresponds to $\lambda = 1$?

- (●) $\theta = \begin{bmatrix} 13.01 \\ 0.91 \end{bmatrix}$
- ( ) $\theta = \begin{bmatrix} 81.47 \\ 12.69 \end{bmatrix}$

Which of the following statements about regularization are true? Check all that apply. 1 point

- [ ] 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.
- [ ] 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.
- [ ] 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).
- [x] Using too large a value of λ can cause your hypothesis to underfit the data.