Chapter 6 Concept Test

Due May 6 at 12pm **Points** 1 **Questions** 4 **Time Limit** None

Instructions

Take this concept test after completing your pre-class preparation for chapter 6 material.

Attempt History

	Attempt	Time	Score
LATEST	Attempt 1	7 minutes	1 out of 1

ⓘ Correct answers are hidden.

Score for this quiz: **1** out of 1
Submitted May 6 at 9:26am
This attempt took 7 minutes.

Question 1 — 0 / 0 pts

The book describes three algorithms for choosing which are the features to include in a machine learning model. Which of these methods is a greedy algorithm (such as “hill climbing”) that does not guarantee an globally optimal solution. Select all that apply.

- [x] Backward Stepwise Selection
- [ ] Best Subset Selection
- [x] Forward Stepwise Selection

Both Forward Stepwise and Backward Stepwise selection are greedy, while Best Subset selection is a brute force algorithm. As long as all of the predictors are independent (there are no interactions between features) then a greedy search is admissible… otherwise a greedy algorithm is not appropriate. For example, if there is an interaction between two or more features which are individually weak predictors then they could together have a better predictive power than a single strong predictor that would have been selected earlier in forward selection. In this way, if only two predictors are allowed, a greedy algorithm could choose the wrong two predictors to use and would not be globally optimal.

Incorrect

Question 2

0 / 0 pts

Which of the following best describes the effects of changing λ (lambda) in a ridge regression or lasso model fit cost function?

- If the model's coefficients are held constant, increasing the value of lambda increases the cost of the fit.
- If the model's coefficients are held constant, increasing the value of lambda does not change the cost of the fit.
- If the model's coefficients are held constant, increasing the value of lambda decreases the cost of the fit.

Because lambda is a multiplier of the squared coefficients (in ridge regression) or the absolute value of the coefficients (in LASSO), the cost function will increase as lambda is increased.

Lambda controls the impact of the coefficients on the cost. Since the coefficient absolute values are used for the lasso, and squared coefficients are used for ridge regression, both values are positive. Increasing lambda when the coefficients are held constant will increase cost. Alternately, if cost is held constant, increasing lambda forces the magnitude of the coefficients to decrease (move closer to zero) in order to maintain the equality of the equation.

Question 3

0 / 0 pts

Suppose you are working on a model which has many features (high dimensional). If the training set accuracy is perfect, but you are unsatisfied with its validation performance, which of the following things could you try to do? Select all that apply.

- [x] Obtain more training data and refit the model

  This is one way to attempt to solve the problem

- [ ] Select a model with more flexibility (example: K-NN with a low number for k)

- [ ] Add higher order polynomial terms of the original features.

- [x] Use a technique to reduce the number of features in the model and refit the model

  This is one way to attempt to solve the problem

The problem you are experiencing is the failure of the model to *generalize* to good performance in validation. Because the training set performance is perfect, you can rule out that the model is not flexible enough to capture the phenomenon. We *don't need a more flexible model* as it won't improve performance. Our model is likely overfitting to the data. To solve an overfitting problem, we can either *increase the amount of data available* to eliminate the chance that the model is learning the noise... or *reduce the number of features being used* - which in turn requires less data for a good model fit.

Question 4

1 / 1 pts

Please answer the following questions in text form. Be specific - wherever possible, include page numbers, filenames, concept names to help your instructor understand what you are referring to:

What was the most confusing aspect of the material you reviewed?

Your Answer:

The process of applying principal components regression was the least clear to me. I understand the point and the result of PCA but the application process in the text was confusing.

Quiz Score: **1** out of 1