Chapter 4 Concept Test

Due Apr 15 at 12pm   **Points** 1   **Questions** 6   **Time Limit** None
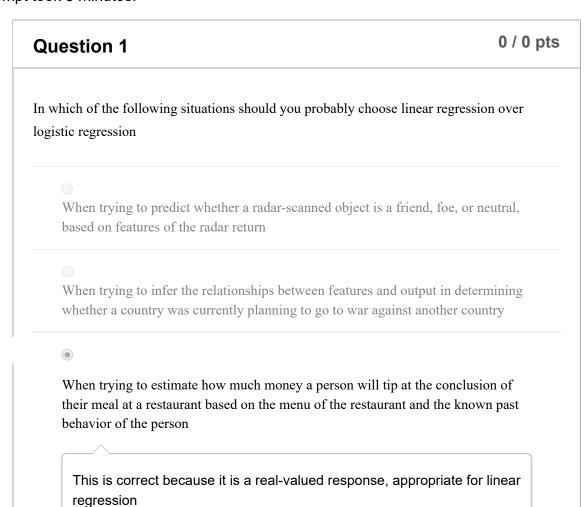
Instructions

Take this concept test after completing your pre-class preparation for chapter 4 material.

Attempt History

	Attempt	Time	Score
LATEST	**Attempt 1**	8 minutes	1 out of 1

Score for this quiz: **1** out of 1
Submitted Apr 15 at 10:09am
This attempt took 8 minutes.

Question 1   0 / 0 pts

In which of the following situations should you probably choose linear regression over logistic regression

○ When trying to predict whether a radar-scanned object is a friend, foe, or neutral, based on features of the radar return

○ When trying to infer the relationships between features and output in determining whether a country was currently planning to go to war against another country

Correct!   ◉ When trying to estimate how much money a person will tip at the conclusion of their meal at a restaurant based on the menu of the restaurant and the known past behavior of the person

> This is correct because it is a real-valued response, appropriate for linear regression

○ When trying to predict whether a person will graduate from college based on their SAT scores

---

Question 2   0 / 0 pts

Which of the following is a situation when you could use Linear Discriminant Analysis (LDA) instead of Logistic Regression?

Use LDA when the number of observations (n) is large

○

'ou Answered

◉

Use LDA when the class distributions are close together (hard to separate)

> When classes are close together, linear regression should work well... LDA may face challenges if the data do not have a gaussian distribution or if there is a large amount of overlap between classes

orrect Answer

○ Use LDA when there are more than 2 classes (example: cat, dog, mouse)

○

Use LDA when the distribution of the predictors/features is non-Gaussian (example: Uniform distribution)

Question 3   0 / 0 pts

Suppose you build a 2-class LDA classifier for detecting a disease based on observations of the patient. Note that **screening tests** are designed to err on the side of caution (rather than wrongly assuming the patient is healthy). You get the following results when you use

"Patient has disease if $\Pr(\text{disease}{=}\text{yes}|X{=}x) > c$", where c is the cutoff probability setting for the test.

		True Status		
		Yes (positive)	No (negative)	Total
predicted	Yes (positive)	50	20	70
	No (negative)	30	900	930
	Total	80	920	1000

What action should you take next if you want to improve the performance of this screening test?

Your Answer:

If we want to improve the performance we can reduce the value of c so that the threshold is lower.

> Set the cutoff probability C lower such that the test predicts more observations are positive.   Ideally in the screening test we do not want to allow people with the disease to be missed (false negative), we must try to get more of the 30 false negatives to be predicted as positive (true positive) by decreasing the c value.

Question 4

0 / 0 pts

When trying to decide whether to use a Quadratic Discriminant Analysis model instead of an LDA model, which of the following situations indicate the use of QDA:

○ There are many observations and the bayes-optimal boundary is a straight line through feature space

Correct!

○ There are many observations and each class has very different feature covariance for its within-class observations

> In this situation, QDA will benefit from the added observations, especially given that each class has a different covariance.

○ There are few observations and each class has different correlations between features of within-class observations

○ There are few observations and the bayes-optimal boundary is a straight line through feature space

---

Question 5   0 / 0 pts

Be specific - wherever possible, include page numbers, filenames, concept names to help your instructor understand what you are referring to:

What was the most SALIENT concept in the material you reviewed for this pre-class assignment? In other words, what really resonated with you on this material?

Your Answer:

Logistic regression resonated the most as I have done some things with it before.

Question 6   1 / 1 pts

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:

Most confusing was how to obtain and utilize all of the statistical attributes for QDA, namely the different covariance matrices.

Quiz Score: **1** out of 1