# C-482 Machine Learning

# Linear Models and Metrics

# Chapter 3

# Homework Questions

Linear Regression

1. A regression analysis relating test scores (*Y* ) to training hours (*X*) produced the following fitted question: ˆ*y* = 25 *−* 0*.*5*x*.

(a) What is the fitted value of the response variable corresponding to *x* = 7?

(b) What is the residual corresponding to the data point with *x* = 3 and *y* = 30?

(c) If *x* increases 3 units, how does ˆ*y* change?

(d) An additional test score is to obtained for a new observation at *x* = 6. Would the test score for the new observation necessarily be 22? Explain.

(e) The error sums of squares (*SSE*) for this model was found to be 7. If there were *n* = 16 observations, provide the best estimate for *σ* 2.

1. Explain the different between the following two equations:

ˆ *Y* = *b*0 + *b*1*X*

*Y* = *β*0 + *β*1*X* + *ϵ.*

1. Given the following predictor matrix X with 2 predictors and sample size of 3,

X =

and the target y,

y =

1. Find the parameters of the linear model β by solving the equation

β = (XT. X(-1))XTy

1. Predict the value of y^ for xi = [ 4 2] xj = [ 1 5]
2. Find R2 for the model developed in a)
3. Briefly explain why one would use lasso penalty and why one would use ridge penalty.
4. Given the following table

Predicted Observed

Cats Dogs

Cats *8 7*

Dogs *5 3*

Compute Accuracy, Recall, Precision and Specificity and f-statistic