MATE-T580: Quiz 7

Name:

Question 1

Which of the following is **NOT** a proper model performance metric for a linear regression:

A.

Mean Squared Error

B.

R-squared

C.

Residual sum of squares

D.

Error rate

Question 2

Which of the following is **NOT** a proper model performance metric for a logistic regression:

A.

Area under the curve

B.

Adjusted R-squared

C.

Error rate

D.

Accuracy

Question 3

Here's a linear regression performed on the mtcars dataset:

```r
data(mtcars)
mod <- lm(mpg ~ hp + factor(cyl), mtcars)
summary(mod)
```

```
##
## Call:
## lm(formula = mpg ~ hp + factor(cyl), data = mtcars)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -4.818 -1.959  0.080  1.627  6.812
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   28.65012    1.58779  18.044  < 2e-16 ***
## hp            -0.02404    0.01541  -1.560  0.12995
## factor(cyl)6  -5.96766    1.63928  -3.640  0.00109 **
## factor(cyl)8  -8.52085    2.32607  -3.663  0.00103 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.146 on 28 degrees of freedom
## Multiple R-squared:  0.7539, Adjusted R-squared:  0.7275
## F-statistic: 28.59 on 3 and 28 DF,  p-value: 1.14e-08
```

Which of the following statements is correct:

A.

Each unit increase in horse power is guaranteed to reduce fuel efficiency by about 0.024

B.

The relationship between fuel efficiency and number of cylinders is not statistically significant

C.

On average, an 8 cylinder car reduces fuel efficiency by 8.52 over a 4 cylinder car, if horse power is held constant

D.

The model above explains about 27% of the variation in fuel efficiency

.

Question 4

Which of the following statements about the Lasso regularization method is **NOT** correct:

A.

The Lasso method can be applied only to the linear regression, not the logistic regression

B.

The hyperparameter λ aids in variable selection by causing the value of some of the regression coefficients to collapse to zero

C.

The Lasso method assumes that the input variables are normalized

D.

Tuning λ involves building different models for different λ values and comparing models' performance using cross validation

Question 5

Which is the main drawback to using a linear or logistic regression:

A.

They are not computationally efficient algorithms

B.

With large number of predictors, it is difficult to include nonlinear terms and interaction terms without subject knowledge expertise

C.

They cannot be applied to cool problems such as detecting faces in images

D.

They always perform poorly compared to a deep neural network or random forests