Assignment 8

The due date for submitting this assignment has passed. As per our records you have not yet submitted this assignment.

1) Regression is used to

A. produce data mining
B. exploratory data mining
C. descriptive data mining
D. explanatory data mining

No, the answer is incorrect.

Acceptable answers:

A. explanatory data mining

1 point

2) The output of a regression algorithm is usually a

A. real variable
B. integer variable
C. character variable
D. string variable

No, the answer is incorrect.

Acceptable answers:

A. real variable

1 point

3) The linear regression algorithm usually produces the least square error between

A. Input value and output value
B. Input value and target value
C. Output value and target value
D. Model parameter and input value

No, the answer is incorrect.

Acceptable answers:

B. Input value and target value

1 point

4) Consider, p, q to be the independent variables and y to be the dependent variable, which of the following represents a linear regression model?

A. \( y = x_1 + x_2 + x_3 + x_4 + x_5 \)
B. \( y = x_1 \cdot x_2 \cdot x_3 \cdot x_4 \cdot x_5 \)
C. \( y = e^{x_1} \cdot e^{x_2} \cdot e^{x_3} \cdot e^{x_4} \cdot e^{x_5} \)
D. \( y = x_1 + x_2 + x_3 \cdot x_4 \cdot x_5 \)

No, the answer is incorrect.

Acceptable answers:

A. \( y = x_1 + x_2 + x_3 + x_4 + x_5 \)

1 point

5) The linear regression model \( y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \) applied to the data in the table shown below. What is the value of the mean squared error function? \( RMSE \) when \( \beta_0, \beta_1, \beta_2 \geq 0? \)

\[
\begin{array}{c|c|c|c|c}
\hline
x_1 & x_2 & y & \hline
-1.5 & 2 & 0.0 & \\
-1.2 & 1 & 0.3 & \\
-0.9 & 0 & 0.6 & \\
-0.6 & -1 & 0.9 & \\
-0.3 & -2 & 1.2 & \\
0 & -3 & 1.5 & \\
0.3 & -2 & 1.8 & \\
0.6 & -1 & 2.1 & \\
0.9 & 0 & 2.4 & \\
1.2 & 1 & 2.7 & \\
1.5 & 2 & 3.0 & \\
\hline
\end{array}
\]

A. 0.06
B. 0.02
C. 0.00
D. 0.01

No, the answer is incorrect.

Acceptable answers:

C. 0.00

1 point

6) The linear regression model \( y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \) applied to the data in the table shown below. What is the optimal regression model, calculated by minimizing mean squared error?

\[
\begin{array}{c|c|c|c|c}
\hline
x_1 & x_2 & y & \hline
-1.5 & 2 & 0.0 & \\
-1.2 & 1 & 0.3 & \\
-0.9 & 0 & 0.6 & \\
-0.6 & -1 & 0.9 & \\
-0.3 & -2 & 1.2 & \\
0 & -3 & 1.5 & \\
0.3 & -2 & 1.8 & \\
0.6 & -1 & 2.1 & \\
0.9 & 0 & 2.4 & \\
1.2 & 1 & 2.7 & \\
1.5 & 2 & 3.0 & \\
\hline
\end{array}
\]

A. 3.10 \( \times \) 2.32
B. 3.10 \( \times \) 1.32
C. 3.00 \( \times \) 0.90
D. 5.07 \( \times \) 0.498

No, the answer is incorrect.

Acceptable answers:

D. 5.07 \( \times \) 0.498

1 point

7) The linear regression model \( y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots \) is to be fit to the data in the table shown below. What is the optimal regression model, calculated by minimizing mean squared error?

\[
\begin{array}{c|c|c|c|c}
\hline
x_1 & x_2 & x_3 & y & \hline
-1.5 & 2 & 3 & 0.0 & \\
-1.2 & 1 & 3 & 0.3 & \\
-0.9 & 0 & 3 & 0.6 & \\
-0.6 & -1 & 3 & 0.9 & \\
-0.3 & -2 & 3 & 1.2 & \\
0 & -3 & 3 & 1.5 & \\
0.3 & -2 & 3 & 1.8 & \\
0.6 & -1 & 3 & 2.1 & \\
0.9 & 0 & 3 & 2.4 & \\
1.2 & 1 & 3 & 2.7 & \\
1.5 & 2 & 3 & 3.0 & \\
\hline
\end{array}
\]

A. \( y = x_1 \times x_2 \times x_3 \)
B. \( y = x_1 + x_2 + x_3 \)
C. \( y = x_1 \cdot x_2 \cdot x_3 \)
D. \( y = x_1 + x_2 + x_3 \cdot x_2 \cdot x_3 \)

No, the answer is incorrect.

Acceptable answers:

B. \( y = x_1 + x_2 + x_3 \)

1 point

8) Accuracy of a linear regression model usually tests

A. low bias and low variance
B. low bias but high variance
C. high bias but low variance
D. high bias and high variance

No, the answer is incorrect.

Acceptable answers:

A. low bias and low variance

1 point

9) A linear regression prediction problem is often not solving

A. Multivariate regression
B. Autoregression
C. Logistic regression
D. Hierarchical regression

No, the answer is incorrect.

Acceptable answers:

D. Hierarchical regression

1 point

10) If you have multiple linear regression models, you would like to choose one of the following:

A. \( \sum x_i \)
B. \( \sum x_i^2 \)
C. \( \sum x_i \cdot y_i \)
D. \( \sum x_i \cdot y_i^2 \)

No, the answer is incorrect.

Acceptable answers:

C. \( \sum x_i \cdot y_i \)

1 point

11) In principal component analysis, the projected lower dimensional space corresponds to

A. subset of the original co-ordinate axes
B. eigenvectors of the data covariance matrix
C. eigenvalues of the data distance matrix
D. orthogonal vectors to the original co-ordinate axes

No, the answer is incorrect.

Acceptable answers:

B. eigenvectors of the data covariance matrix

1 point