

Exam Sample

Choose the best answer

1-Logistic Regression

Logistic regression is used to predict _____ valued output?

- A) Continuous
- B) Categorical
- C) Both

2- In binary logistic regression:

- A. The dependent variable is continuous.
- B. The dependent variable is divided into two equal subcategories.
- C. The dependent variable consists of two categories.
- D. There is no dependent variable.

3- Logistic regression is used when you want to:

- A. Predict a discrete variable from continuous or discrete variables.
- B. Predict a continuous variable from discrete variables.
- C. Predict any categorical variable from several other categorical variables.
- D. Predict a continuous variable from discrete or continuous variables.

4- A residual is defined as

- A. $Y - Y'$
- B. Error Sum of Square
- C. Regression Sum of Squares

5- The line described by the regression equation attempts to

- A. pass through as many points as possible.
- B. pass through as few points as possible.
- C. minimize the number of points it touches.
- D. minimize the squared distance from the points.

6- A regression model in which more than one independent Features is used to predict the dependent variable is called

- A- a simple linear regression model
- B- a multiple regression model
- C- an independent model
- D- none of the above

7- Suppose, you got a situation where you find that your linear regression model is under fitting the data. In such situation which of the following options would you consider?

- 1- I will add more variables
- 2- I will start introducing polynomial degree variables
- 3- I will remove some variables

- A)1 and 2
- B)2 and 3
- C)1 and 3
- D)1, 2 and 3

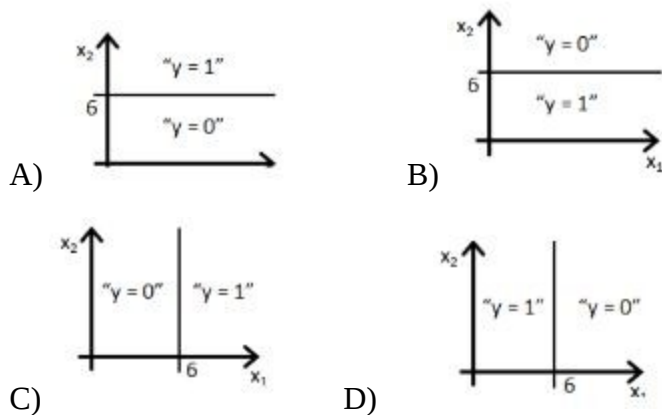
8 -Suppose you train a logistic regression classifier and your hypothesis function H is

$$h_{\theta}(x) = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2)$$

Where

$$\theta_0 = 6, \theta_1 = 0, \theta_2 = -1.$$

Which of the following figure will represent the decision boundary as given by above classifier?



Question: Theoretical Concepts

- 1- Explain the principle of the gradient descent algorithm. Accompany your explanation pseudo-code.
- 2- Describe the Concept of K-cross Validation

Question: Problems

Suppose we have been given training data consisting of the Tumor size(T) of a patient indicating whether the patient has cancer/not cancer in the diagnosis. Using this data we train a logistic regression model, which has the following parameters :

$$\theta_0 = -8.75, \theta_1 = 0.25$$

Estimate the probability for a patient with T=35 to be diagnosed as cancer