# Week 8 : Assignment 8

**The due date for submitting this assignment has passed.**

**Due on 2023-03-22, 23:59 IST.**

## Assignment submitted on 2023-03-19, 12:54 IST

1) Which of the following methods do we use to best fit the data in Logistic Regression? *1 point*

- Least Square Error
- Maximum Likelihood
- Jaccard distance
- All of these

Yes, the answer is correct.
Score: 1
Accepted Answers:
*Maximum Likelihood*

2) Which of the following evaluation metrics can not be applied in case of logistic regression output to compare with target? *1 point*

- AUC-ROC
- Accuracy
- Logloss
- Mean-Squared-Error

Yes, the answer is correct.
Score: 1
Accepted Answers:
*Mean-Squared-Error*

3) Let f(x) denote the logistic function. The range of f(x) for any real value of x is *1 point*

- (0,1)
- (-1 , 1)
- All positive integers
- All negative integers

Yes, the answer is correct.
Score: 1
Accepted Answers:
*(0,1)*

4) Which of the following option is true? *1 point*

- Linear Regression errors values has to be normally distributed but in case of Logistic Regression it is not the case
- Logistic Regression errors values has to be normally distributed but in case of Linear Regression it is not the case
- Both Linear Regression and Logistic Regression error values have to be normally distributed
- Both Linear Regression and Logistic Regression error values have not to be normally distributed

Yes, the answer is correct.
Score: 1
Accepted Answers:
*Linear Regression errors values has to be normally distributed but in case of Logistic Regression it is not the case*

5) For the figure given below, which decision boundary is overfitting the training data? *1 point*

https://drive.google.com/file/d/1LWseSSMlbeirh0GjSGwnEUdYOyeR_ysu/view?usp=share_link

- A
- B
- C
- None of these

Yes, the answer is correct.
Score: 1
Accepted Answers:
*C*

6) Select the correct alternatives from the following based on the figure *1 point*

https://drive.google.com/file/d/1LWseSSMlbeirh0GjSGwnEUdYOyeR_ysu/view?usp=share_link

1. The training error in first plot is maximum as compared to second and third plot.
2. The best model for this regression problem is the last (third) plot because it has minimum training error (zero).
3. The second model is more robust than first and third because it will perform best on unseen data.
4. The third model is overfitting more as compared to first and second.
5. All will perform same because we have not seen the testing data.

- 1 and 3
- 1 and 4
- 1,3 and 4
- 5

Yes, the answer is correct.
Score: 1
Accepted Answers:
*1,3 and 4*

7) For categorical data with 'n' categories, the number of dummy variables will be_____ *1 point*

- n
- n - 1

- n + 1
- 2n

Yes, the answer is correct.
Score: 1
Accepted Answers:
*n - 1*

8) In binary logistic regression, ***1 point***

- The dependent variable is continuous
- The dependent variable is divided into two equal subcategories
- (•) The dependent variable consists of two categories
- There is no dependent variable

Yes, the answer is correct.
Score: 1
Accepted Answers:
*The dependent variable consists of two categories*

9) If the number of False negatives is 5 and number of True Positives is 20, the value of recall will be equal to _____ ***1 point***

- 0.2
- 0.6
- (•) 0.8
- 0.3

Yes, the answer is correct.
Score: 1
Accepted Answers:
*0.8*

10) If the precision is 0.6 and the recall value is 0.4, the value of f-measure will be ***1 point***

- (•) 0.48
- 1
- 0.24
- None of these

Yes, the answer is correct.
Score: 1
Accepted Answers:
*0.48*