# Foundation of ML MDS Quiz - 1

[redacted]

Your email will be recorded when you submit this form

## Foundations of ML

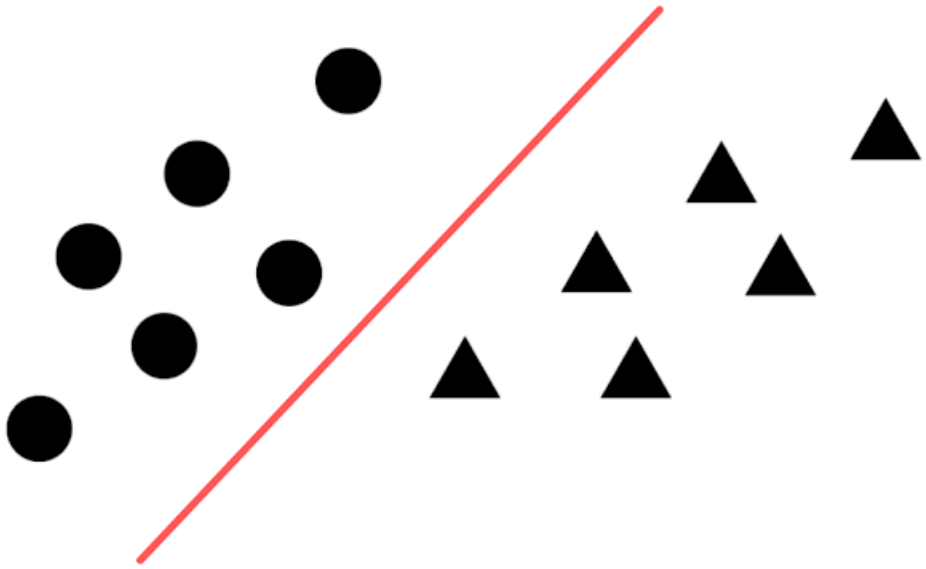
Suppose X1, . . . , Xn constitute a sample from a uniform distribution on (0, θ), where θ is unknown. What is the maximum likelihood estimator of θ ?

- ○ mean(X1, . . . , Xn )
- ○ min(X1, . . . , Xn )
- ○ max(X1, . . . , Xn )
- ○ median(X1, . . . , Xn )

True or False : "In the ML hackathon, two teams are trying to solve the same logistic regression problem for a dataset. Team 1 claims that their initialization point will lead to a much better optimum than Team 2's initialization point."

- ○ True
- ○ False

[True/False] ] Consider the following data , with positive examples labeled as circles and negative examples labeled as triangles. Linear discriminant analysis (LDA) projects the data points to the red line in this figure.



- ◯ True
- ◯ False

Suppose you were interviewed for a technical role. 50% of the people who sat for the first interview received the call for second interview. 95% of the people who got a call for second interview felt good about their first interview. 75% of people who did not receive a second call, also felt good about their first interview. If you felt good after your first interview, what is the probability that you will receive a second interview call?

- ◯ 66%
- ◯ 56%
- ◯ 75%
- ◯ 85%

Suppose that you have a dataset D1 and you design a linear regression model of degree 3 polynomial and you found that the training and testing error is “0” or in another terms it perfectly fits the data. What will happen when you fit degree 2 polynomial in linear regression?

- ◯ It is high chances that degree 2 polynomial will over fit the data
- ◯ It is high chances that degree 2 polynomial will under fit the data
- ◯ Can't say
- ◯ None of these

[Marks : 4] What is the least squares solution to the problem with the following data (assume the line passes through the origin):

$$X = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & 1 \end{bmatrix} Y = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$\begin{bmatrix} \frac{1}{3} & -\frac{1}{3} \end{bmatrix}$

Option 1

$\begin{bmatrix} -\frac{1}{3} & \frac{1}{3} \end{bmatrix}$

Option 2

$\begin{bmatrix} \frac{1}{3} & \frac{1}{3} \end{bmatrix}$

Option 3

$\begin{bmatrix} -\frac{1}{3} & -\frac{1}{3} \end{bmatrix}$

Option 4

Consider the following statements:

i) As the number of data points grows to infinity, the MAP estimate approaches the MLE estimate for all possible priors. In other words, given enough data, the choice of prior is irrelevant.

ii) MLE is equivalent to MAP when the prior is a normal distribution with a standard deviation of infinity.

- ◯ (i) is True and (ii) is False
- ◯ (i) is False and (ii) is True
- ◯ Both are True
- ◯ Both are False

[True/False] Fisher linear discriminant maximizes wS1w, where S1 is

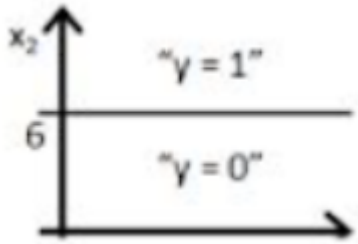
$$\mathbf{S}_1 = \sum_t \left(\mathbf{x}^t - \mathbf{m}_1\right)\left(\mathbf{x}^t - \mathbf{m}_1\right)^T r^t$$

- ◯ True
- ◯ False

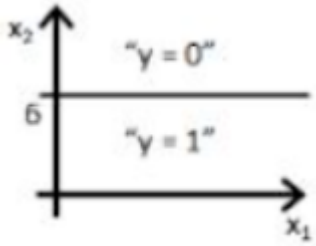
Assume you have read a paper or a blog about a machine learning algorithm and the author makes some claims about the algorithm. Which one of the following is the most reasonable claim to accept ?

- "My algorithm is better than previous ones. Look at the training error rates!"
- "My algorithm is better than previous ones. Look at the test error rates!  Results are reported for best value of hyperparameter on  train data"
- "My algorithm is better than previous ones. Look at the test error rates! Results are reported results for best value of hyperparameter, chosen with 10-fold cross validation."
- "My algorithm is better than previous ones. Look at the train error rates!  Results are reported results for best value of hyperparameter on test data"

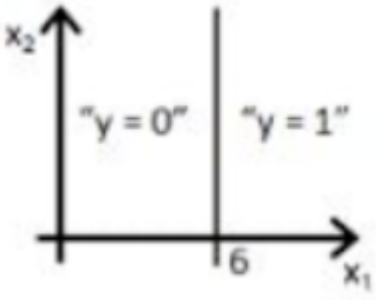
[Marks: 4] Suppose you train a logistic regression classifier and your hypothesis function H is h(x1,x2) = g(a + b x1 + c x2 ) where a=6, b=0, c = -1. Which of the following figure will represent the decision boundary as given by above classifier?



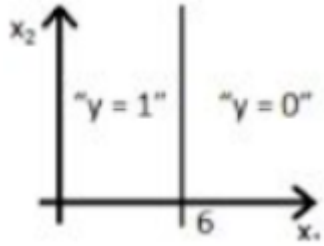

Option 1




Option 2

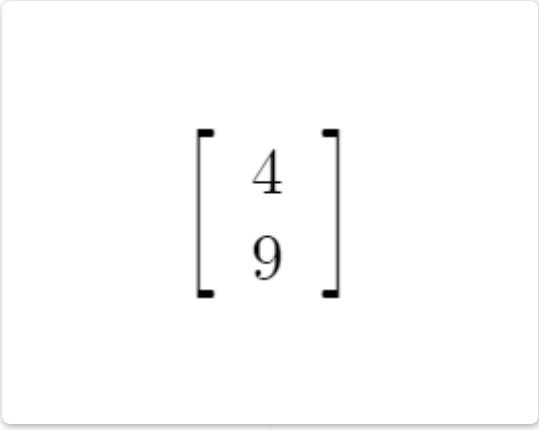



Option 3

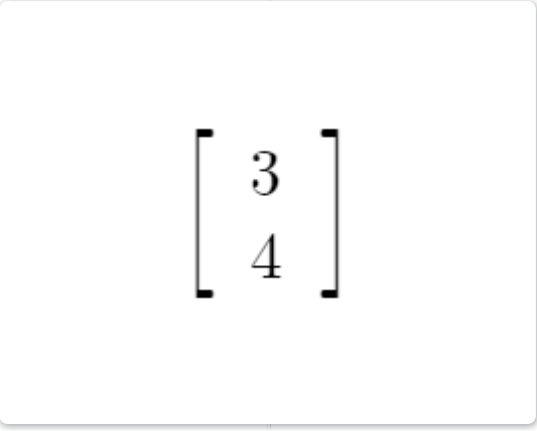



Option 4

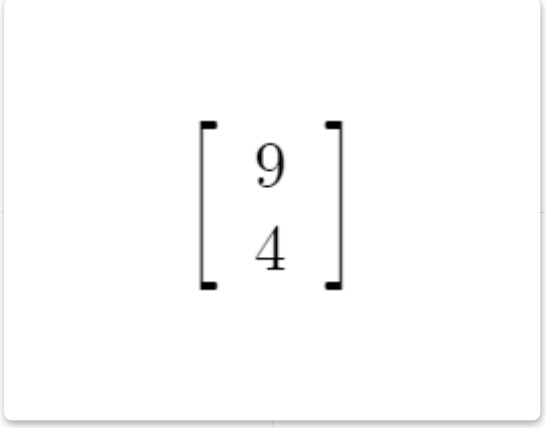
[Marks : 4] Write down the equations for the line y = mx+c to go through y =7 at x = -1, y=7 at x=1 and y=21 at x=2. Find the least squares solution(c,m).



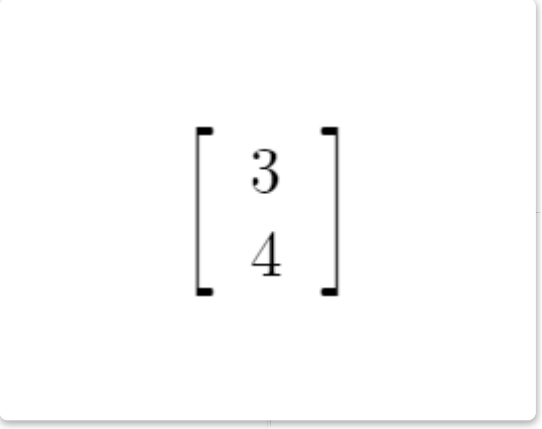
Option 1



Option 2

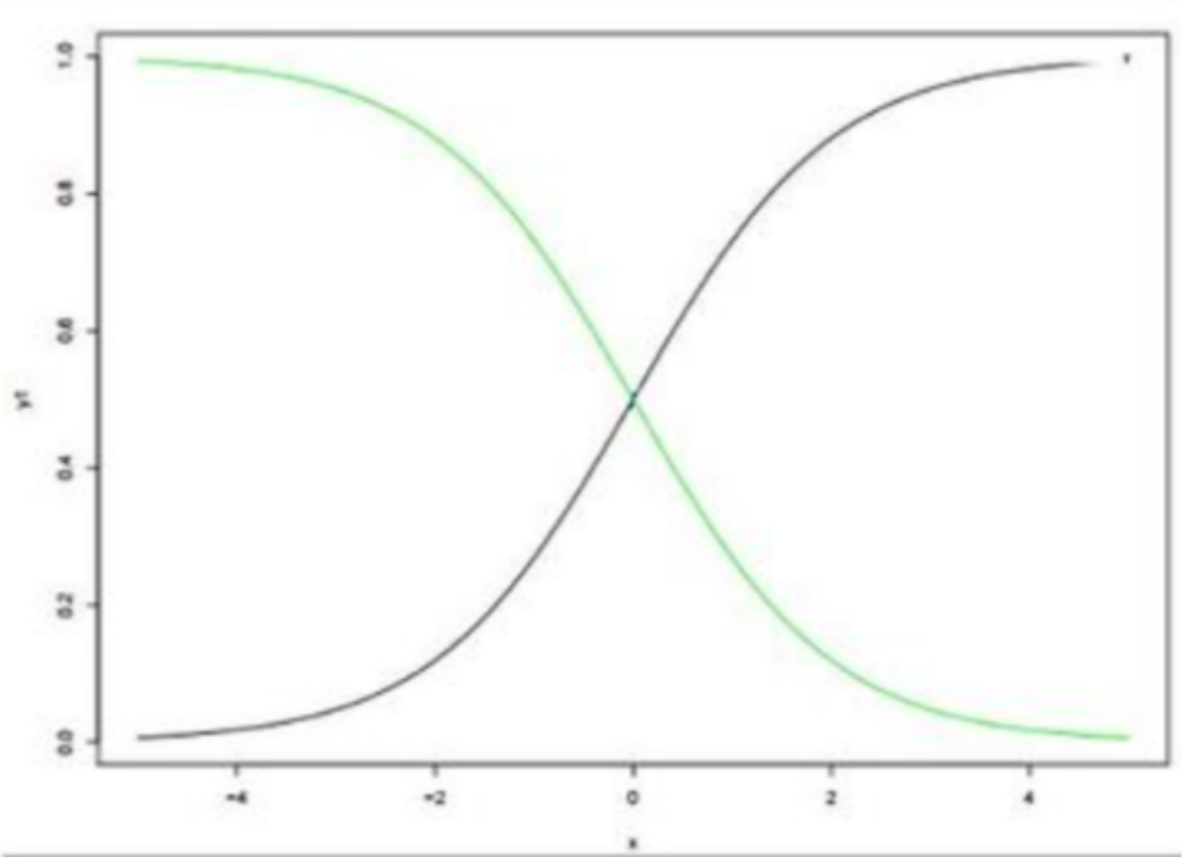


Option 3



Option 4

[Marks : 4] Below are two different logistic models with different values for 'm' and 'c'. Which of the following statement is true about 'm' and 'c' values if two logistics models(green,black)? Consider y = mx+c, 'c' is the intercept and 'm' is coefficient

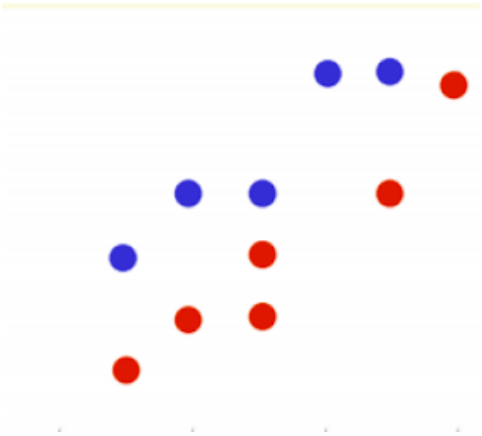


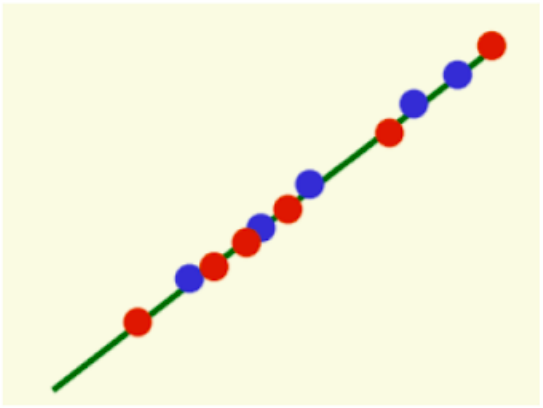

- 'm' for Green is greater than black
- 'm' for green is lower than black
- 'm' for both models is same
- cant say

True or False : "Discriminant analysis concentrate on learning just the decision boundary between classes."

- True
- False

[True/False] Consider the following data , with positive examples labeled as blue and negative examples labeled as red. The figure on the right side represents the Fisher linear discriminant projection.





- ◯ True
- ◯ False

[Marks : 4] Consider a survey where people are asked if they buy ipad or galaxy tab. Assume 8 out of 10 said ipad. Assuming a Gamma (a,b) prior distribution with a= 3 and b=4 over probability p of people likely to buy ipad the MAP estimate of p ?

- ◯ 2/3
- ◯ 3/4
- ◯ 4/5
- ◯ 11/17

k-fold cross-validation is a model selection method used to choose hyperparameters when training machine learning models. Which of the following is correct about the time complexity of k-fold cross-validation?

- ○ Is linear in k
- ○ Is quadratic in k
- ○ Is cubic in k
- ○ Is exponential in k

Consider the statements:

i) Linear regression with L1 regularization is useful for feature selection.
ii) Linear regression with L2 regularization reduces variance
iii) Bayesian linear regression never overfits on the data.
iv) Bayesian linear regression provides confidence bounds.

which among the following is correct:

- ○ Only ii
- ○ i and ii only
- ○ ii and iv only
- ○ All of the above

Imagine, you are working with a news agency and you want to develop a machine learning algorithm which predicts the number of views on the newly published articles. Which of the following evaluation metric would you choose in that case?
1. mean square error 2. accuracy 3. recall

- only 1
- only 2
- only 3
- 1 and 3
- 2 and 3
- 1 and 2

In logistic regression, log-odds ratio is a

- logistic function of x
- linear function of x
- quadratic function of x
- probit function of x

Consider the following statement:
i) Linear regression with L2 regularization does not have a closed form solution.
ii) Linear regression with L1 regularization has closed form solution.
iii) Linear regression without any regularization will always have a unique solution.
iv) Linear regression with L2 regularization will not have unique solution.
which among the following is correct

- ○ i
- ○ ii
- ○ iii
- ○ iv
- ○ i and iii
- ○ None of the above

Back Submit Clear form