Foundation of ML Quiz - 1

Maximum duration is 30 mts. You can resubmit until 8:30 PM, but not beyond 8:30 PM. No negative marking.

The respondent's email (**cs18btech11001@iith.ac.in**) was recorded on submission of this form.

Course ID (FoML course ID you have registered in AIMS : CS5590, AI5000, SM5000, AI2000) *

- (●) CS5590
- ( ) AI5000
- ( ) SM5000
- ( ) AI2000

Program (PhD, MTech, MDS, BTech) *

- (●) BTech
- ( ) MTech
- ( ) MDS
- ( ) PhD

Name *

Abburi Venkata Sai Mahesh

Roll no *

CS18BTECH11001

Foundations of ML

True or False : "Discriminative methods in classication approaches concentrate on learning just the decision boundary between classes."

- (●) True
- ( ) False

Consider modeling rain prediction (if it rains or not) using binary random variable X. X takes value 1 if it rains and 0 if it does not rain. X is Bernoulli distributed with parameter q (probability it rains). You have observed one day data where it did not rain (X=0). Assume a uniform prior distribution over q, compute posterior over q and use it to compute predictive probability that it will rain the next day using Bayesian inference. The predictive probability according to Bayesian inference is:

- ( ) 0
- ( ) 1
- (●) 1/3
- ( ) 1/2

What are the natural parameters of the following distribution over x with parameters a and y.

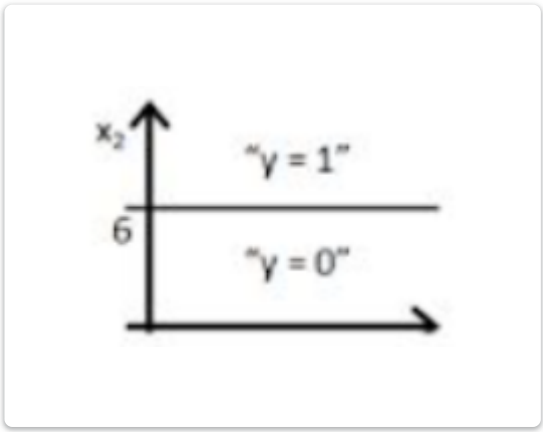
$$\frac{ay^a}{x^{a+1}}$$

- a+1
- -a-1
- a
- log x

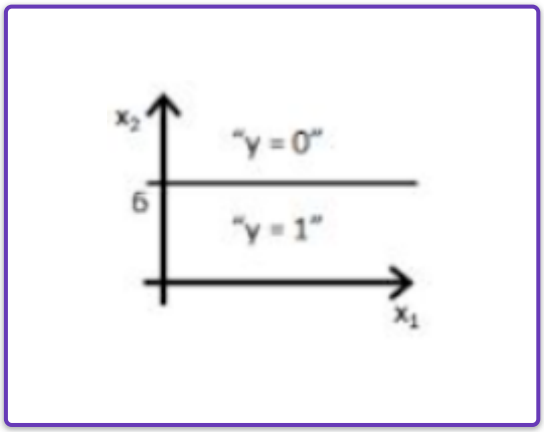
Correct answer

- -a-1

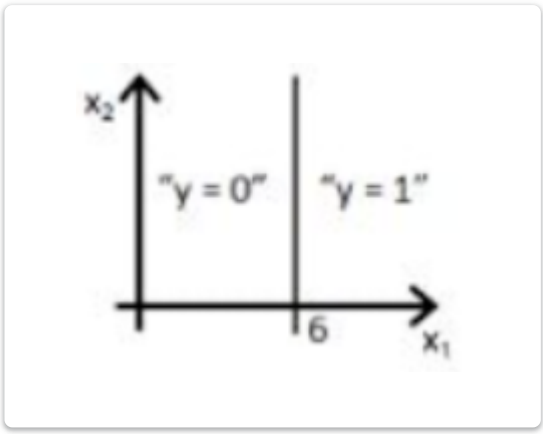
Suppose you train a logistic regression classifier and your hypothesis function H is h(x1,x2) = g(a + b x1 + c x2) where a=6, b=0, c = -1. Which of the following figure will represent the decision boundary as given by above classifier?



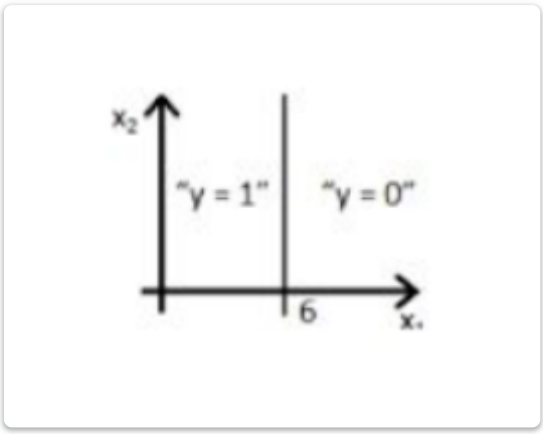
Option 1



Option 2



Option 3



Option 4

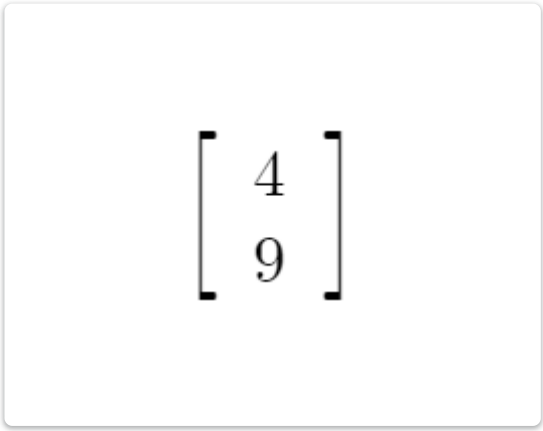
Consider the following statement:
i) Linear regression with L2 regularization does not have a closed form solution.
ii) Linear regression with L1 regularization has closed form solution.
iii) Linear regression without any regularization will always have a unique solution.
iv) Linear regression with L2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

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

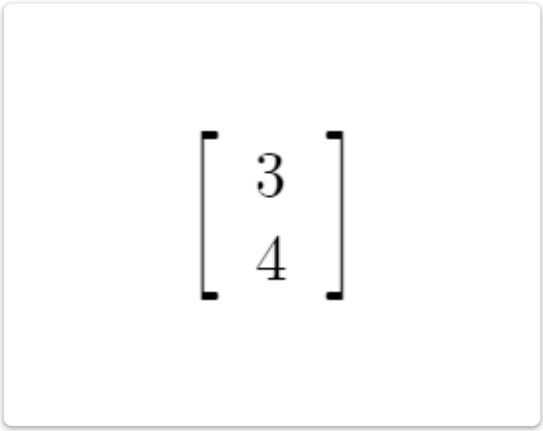
Correct answer

- ◉ None of the above

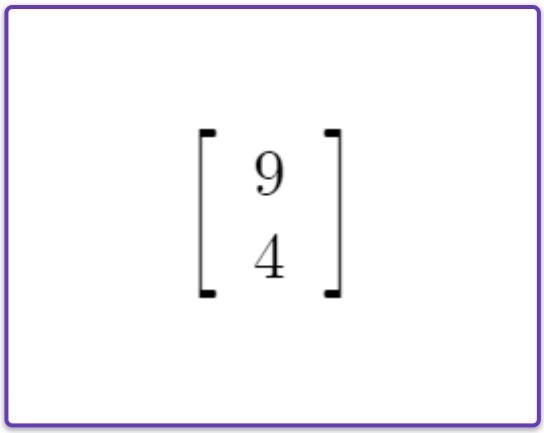
Write down the equations for the line y = mx+c to go through y =7 at x = -1, y=7 at x=1 and y=21 at x=2. Find the least squares solution(c,m).



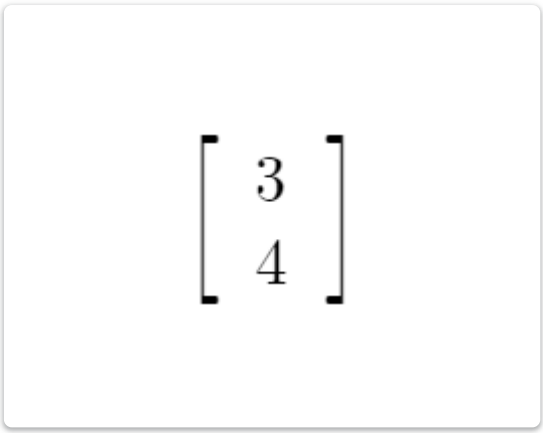
Option 1



Option 2



Option 3



Option 4

Suppose you are only allowed to use binary classifiers to solve a multi-class classication problem: for example consider logistic regression with logistic likelihood. Assuming you are using the one-vs-one approach where you train a classifier to classify every pair of classes, what is the optimal number of classifiers you will need to solve the multi-class problem with 6 classes ?

- ○ 25
- ◉ 15
- ○ 18
- ○ 12

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

Correct answer

- ◉ It is high chances that degree 2 polynomial will under fit the data

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

Let's say, you are working with multivalued categorical feature(s) and you want to use a Naive Bayes model to do classification using these features. Which of the following distributions would you choose to model class conditional density.

- (●) Multinoulli
- ( ) Gaussian
- ( ) Bernoulli
- ( ) exponential

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

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

Correct answer

- (●) All of the above

Find the least squares solution to the problem with data:

$$A = \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & 1 \end{bmatrix} b = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

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

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

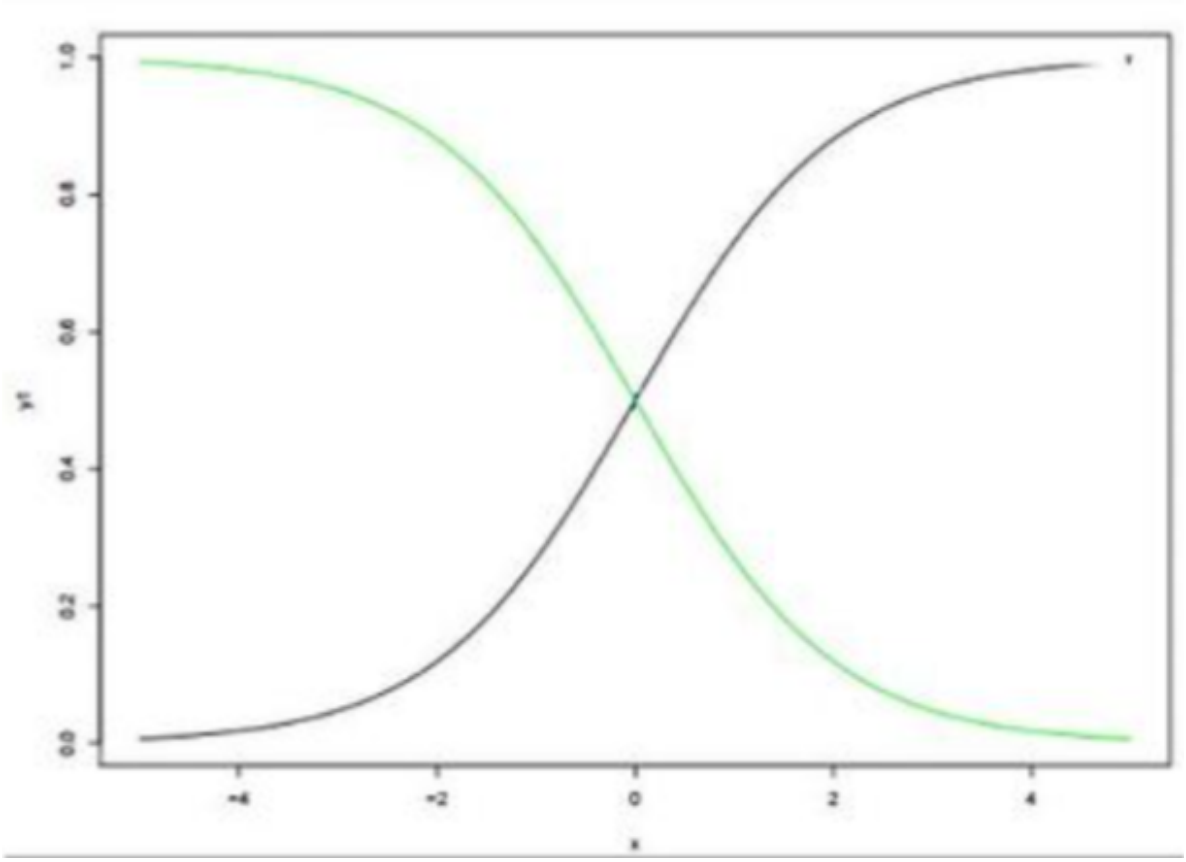
In logistic regression, log-odds ratio is a

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

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%

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

A key characteristic of any machine learning model is its bias and variance. Now, consider you averaged multiple logistic regression models to get the prediction output. Which one of the following is the most expected outcome of this when compared to the predictions of a single logistic regression model?

- ○ Decease bias
- ○ Increase bias
- ◉ Decrease variance
- ○ Increase variance

No correct answers

The difference(s) between generative models and discriminative models is(are):

- ○ Discriminative models capture the joint distribution between features and class labels;
- ○ Generative models assume conditional independence among features;
- ◉ Generative models can effectively explore unlabeled data;
- ○ Discriminative models provide more flexibility in introducing features.

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)

No correct answers

True or False : "Decision boundary obtained using a logistic regression model can be obtained using a generative model with class-conditional Gaussian distribution"

- (●) True
- ( ) False

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

Correct answer

- (●) Is linear in k

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 hyperparameter value 1:7894"
- ◉ "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"

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

This form was created inside of IIT Hyderabad.

Google Forms