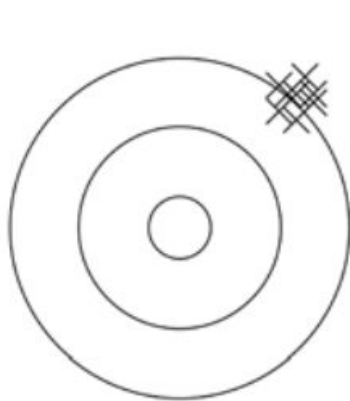


Introduction to Machine Learning

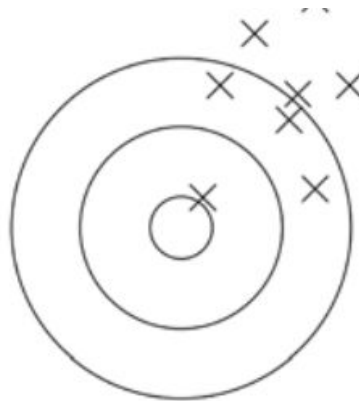
Quiz - Saptahaant Shikshak Prashikshan (SSP)

There are 6 questions in the quiz. Please answer all the questions. Each question carries 5 marks and there is no negative marking.

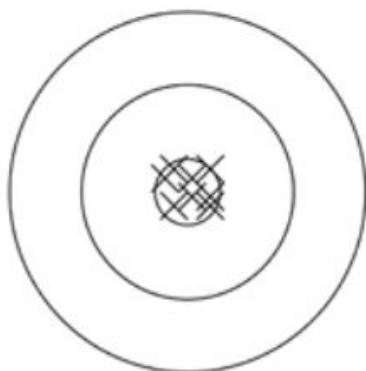
- 1) You are given a sequence with 16 observations, 0 0 0 1 2 1 1 1 1 2 1 1 0 1 2 2. You are told that you have to design an HMM with 4 hidden states. Calculate the total number of parameters your HMM will have. Write the dimensions for the transition matrix, emission matrix and the initial state vector.
- 2) Given below are results for the international shooting competition from 4 participants namely a,b,c and d. In a shooting competition each participant have to shoot the bullets onto the circular boards. Ideally all the bullets should fall within the innermost circle in the board. With your recently acquired skills from machine learning, please identify the variance and bias (from the ideal outcome) shown by each of the participants based on the bullet marks on the boards. A participant may either show high or low bias and high or low variance. For example, participant c has got all the shoot marks within the innermost circle. So we can say that the participant shows low variance and low bias. Similarly, Fill in the variance and bias values for each of the participants. Please note that you have to mention whether the asked value is high or low



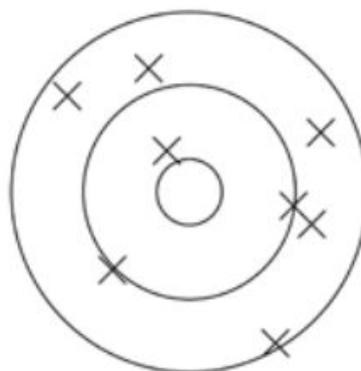
a) Variance: ? ; Bias: ?



b) Variance: ? ; Bias: ?



c) Variance: Low ; Bias: Low



d) Variance: ? ; Bias: ?

Introduction to Machine Learning

Quiz - Saptahaant Shikshak Prashikshan (SSP)

- 3) Given the following confusion matrix for a classifier you have built, calculate the Precision, Recall and Accuracy for your classifier.

	Predicted : No	Predicted: Yes
Actual : No	50	10
Actual : Yes	5	100

- 4) Fill in the blanks with suitable answers for the following questions

- a) _____ is a way to decrease the variance of your prediction by generating additional data for training from your original dataset using combinations with repetitions to produce multisets of the same cardinality/size as your original data. By increasing the size of your training set you can't improve the model predictive force, but just decrease the variance, narrowly tuning the prediction to expected outcome.
- b) _____ is a two-step approach, where one first uses subsets of the original data to produce a series of averagely performing models and then "boosts" their performance by combining them together using a particular cost function, say majority vote.
- c) _____ is a technique for assessing how the results of a statistical analysis will generalize to an independent data set. A single round involves partitioning a sample of data into complementary subsets, performing the analysis on one subset, and validating the analysis on the other subset. To reduce variability, multiple rounds are performed using different partitions, and the results are combined (e.g. averaged) over the rounds to estimate a final predictive model.
- d) You are given a task to predict the diameter of a set of the cylindrical containers. The features and diameters for different cylinder samples are already provided to you. You are also said that the predicted diameters should be a real number with a precision 2 decimal points within the range of [0,12.5]. Which of the following method(s) will be used for prediction of such a model? Please note that more than one option may be correct. You will have to give all the correct options. There will be no partial scores.
- i) Linear regression
 - ii) Logistic Regression
 - iii) K-Means Clustering
 - iv) Naive Bayes.

Introduction to Machine Learning

Quiz - Saptahaant Shikshak Prashikshan (SSP)

- 5) You have to design a linear regression model where the cost function (J) is given below. θ is the parameters of the model. You have to optimize the cost function with respect to the parameters to get the final parameter values.

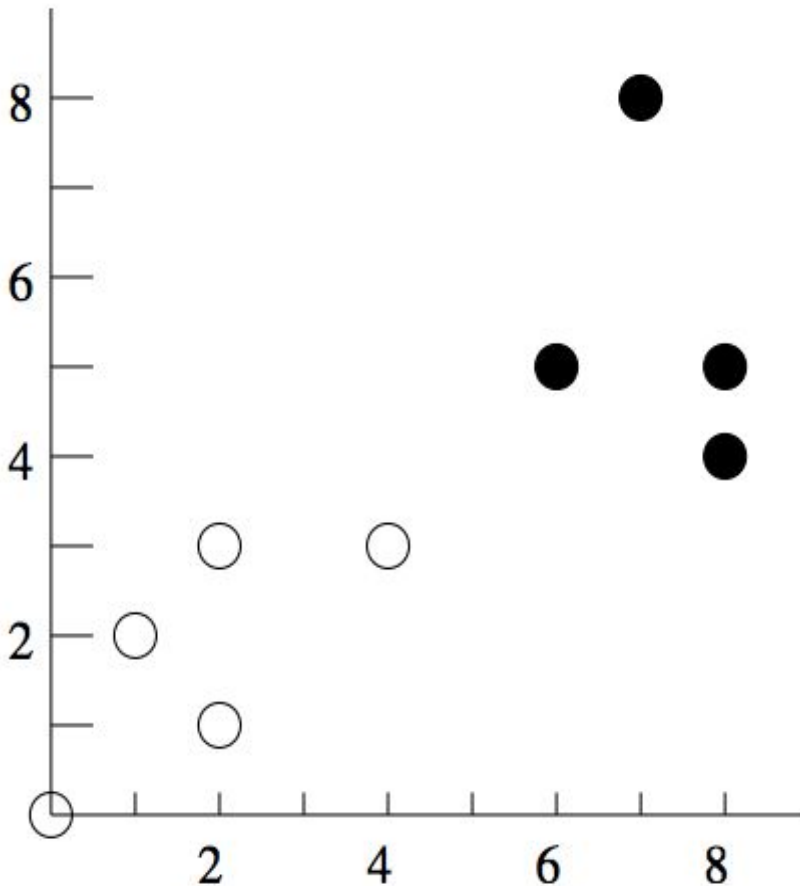
$$J(\theta) = \frac{1}{m} \sum_{i=1}^m (\theta^T x_i - y_i)^2$$

Write the gradient descent equation for optimizing the cost function.

Consider the gradient descent update equation:

$$\theta_{n+1} := \theta_n - \alpha \nabla J(\theta_n)$$

where α is the learning rate



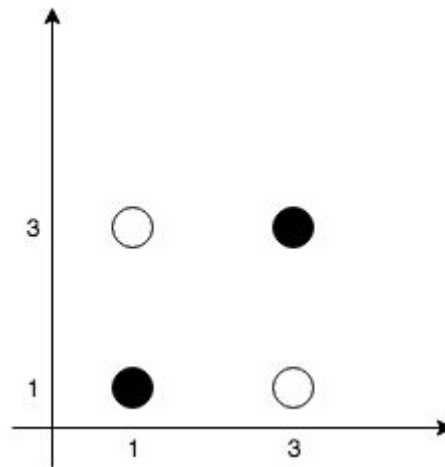
- 6) Answer the following questions:

- (a) Consider a binary classification problem where the data in the graph are taken to be training data. Black and white dotted points correspond to different labels. You have

Introduction to Machine Learning

Quiz - Saptahaant Shikshak Prashikshan (SSP)

to model Support Vector Machine (SVM) for the binary classification. Which of the instances should identified as support vectors? (write in terms of coordinate values)



(b) Consider a binary classification problem where the data given in the graph are taken to be training data. Black and White dotted points correspond to two different labels.

Which of the following kernel(s) can classify the data while using Support Vector Machine? Please note that more than one option may be correct. You will have to give all the correct options. There will be no partial scores.

- i) Linear Kernel
- ii) Polynomial Kernel of degree 2
- iii) Gaussian Kernel
- iv) None of these