MAULANA ABUL KALAM AZAD UNIVERSITY OF TECHNOLOGY, WEST BENGAL

Paper Code : PEC-CS701E Machine learning

Time Allotted : 3 Hours *Full Marks :70*

The Figures in the margin indicate full marks.
Candidate are required to give their answers in their own words as far as practicable

Group-A (Very Short Answer Type Question)

1. Answer any ten of the following : [1 x 10 = 10]

(I) _____ Is a classification algorithm used to assign observations to a discrete set of classes.

(II) The number of nodes in the input layer is 10 and the hidden layer is 5. The maximum number of connections from the input layer to the hidden layer are

(III) True or False: Hierarchical clustering is slower than non-hierarchical clustering?

(IV) True or False: Ensemble learning can only be applied to supervised learning methods.

(V) A collection of individual models that learn to predict a target by combining their strengths and avoiding the weaknesses of each is called _____

(VI) Semi-supervised learning algorithm deals with which types of data

(VII) In an election, N candidates are competing against each other and people are voting for either of the candidates. Voters don't communicate with each other while casting their votes. Which of the following ensemble method works similar to above-discussed election procedure?

(VIII) A feature F1 can take certain value: A, B, C, D, E, & F and represents grade of students from a college.Feature F1 is an example of dependent variable.

(IX) Imagine a Newly-Born starts to learn walking. It will try to find a suitable policy to learn walking after repeated falling and getting up. Specify what type of machine learning is best suited?

(X) The selling price of a house depends on many factors. For example, it depends on the number of bedrooms, number of kitchen, number of bathrooms, the year the house was built, and the square footage of the lot. Given these factors, predicting the selling price of the house is an example of which type of linear regression.

(XI) Targeted marketing, Recommended Systems, and Customer Segmentation are applications in which algorithm?

(XII) The _____ is the difference between a sample statistic used to estimate a population parameter and the actual but unknown value of the parameter.

Group-B (Short Answer Type Question)

Answer any three of the following [5 x 3 = 15]

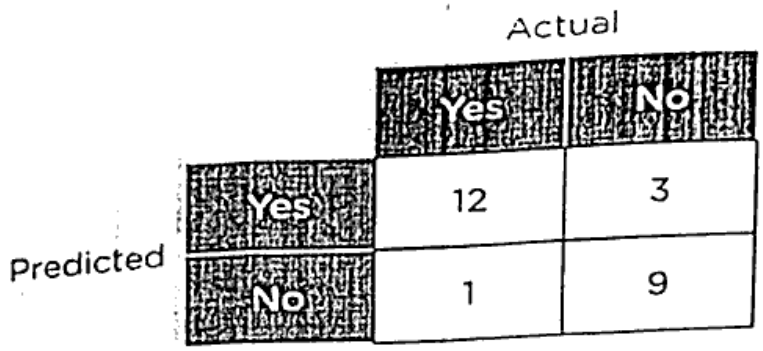
2. Explain Matrix Factorization and where it is uesd. [5]
3. Why ensemble learning is used? What is the general principle of an ensemble method and what is bagging and boosting in ensemble method? [5]
4. Explain the Difference Between Classification and Regression? [5]
5. Compare K-means and KNN Algorithms. [5]
6. How do we decide the value of "K" in KNN algorithm? Why is the odd value of "K" preferable in KNN algorithm? [5]

Group-C (Long Answer Type Question)

Answer any three of the following [15 x 3 = 45]

7. (a) Discuss the different types of Machine Learning? [6]
   (b) What are parametric and non-parametric model? [5]
   (c) How is machine learning related to AI? [4]
8. (a) Explain Generative Mixture model [4]
   (b) With a proper diagram explain the steps of a genarative mixture model [5]
   (c) Write down the steps of PCA (Principal Component Analysis) [6]
9. (a) Explain the Confusion Matrix with Respect to Machine Learning Algorithms with an suitable example [5]

(b) Calculate the accuracy percentage for the given Confusion Matrix. [5]



Confusion Matrix

(c) Explain True Positive, True Negative, False Positive, and False Negative in Confusion Matrix with an example. [5]

10. (a) Explain the three techniques under supervised feature Selection [6]

(b) Explain the benefits of using feature selection in machine learning [4]

(c) Explain the curse of dimensionality [5]

11. (a) What is Artificial Intelligence and why do we need it? [5]

(b) What is Deep Learning, and give some of its example that is used in real-world? [4]

(c) Differentiate between Artificial intelligence, Machine Learning, and Deep Learning [6]

*** END OF PAPER ***