MAULANA ABUL KALAM AZAD UNIVERSITY OF TECHNOLOGY, WEST BENGAL

Paper Code : OE-ME 701H 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) What are two techniques of Machine Learning?
(ii) What are the areas in robotics and information processing where sequential prediction problem arises?
(iii) What is PAC Learning?
(iv) What according to you, is more important between model accuracy and model performance?
(v) What is test set?
(vi) What are the advantages of Naive Bayes?
(vii) What do you understand by Cluster Sampling?
(viii) What do you understand by the F1 score?
(ix) What is a Random Forest?
(x) What could be the issue when the beta value for a certain variable varies way too much in each subset when regression is run on different subsets of the given dataset?
(xi) What is the Principle Component Analysis?
(xii) Why is rotation of components so important in Principle Component Analysis (PCA)?

Group-B (Short Answer Type Question)

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

2. In what areas Pattern Recognition is used? [5]
3. What are the different methods for Sequential Supervised Learning? [5]
4. How does Machine Learning differ from Deep Learning? [5]
5. What do you know about Bayesian Networks? [5]
6. What are the components of relational evaluation techniques? [5]

Group-C (Long Answer Type Question)

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

7. (a) What is Linear Regression? [5]
   (b) Differentiate between regression and classification. [3]
   (c) What is target imbalance? How do we fix it? A scenario where you have performed target imbalance on data. Which metrics and algorithms do you find suitable to input this data onto? [7]
8. (a) How would you handle an imbalanced dataset? [8]
   (b) Mention some of the EDA Techniques? [7]
9. (a) Is it possible to test for the probability of improving model accuracy without cross-validation techniques? If yes, please explain. [5]
   (b) How can we use a dataset without the target variable into supervised learning algorithms? [3]
   (c) State the limitations of Fixed Basis Function. [7]
10. (a) Define and explain the concept of Inductive Bias with some examples. [5]
    (b) Explain how a Naive Bayes Classifier works. [5]
    (c) List the advantages and limitations of the Temporal Difference Learning Method. [5]
11. (a) A data set is given to you about utilities fraud detection. You have built aclassifier model and achieved a performance score of 98.5%. Is this a goodmodel? If yes, justify. If not, what can you do about it? [8]

(b) Explain the handling of missing or corrupted values in the given dataset. [4]

(c) What is Time series? [3]

*** END OF PAPER ***