



BRAINWARE UNIVERSITY

Term End Examination 2023

Programme – B.Tech.(CSE)-2019/B.Tech.(CSE)-2020

Course Name – Machine Learning

Course Code - PEC-601C

(Semester VI)

LIBRARY
Brainware University
Barasat, Kolkata -700125

Full Marks : 60

Time : 2:30 Hours

[The figure in the margin indicates full marks. Candidates are required to give their answers in their own words as far as practicable.]

Group-A

(Multiple Choice Type Question)

1 x 15=15

1. Choose the correct alternative from the following :

- (i) define top-down parser
- | | |
|--|--|
| a) Begins by hypothesizing a sentence (the symbol S) and successively predicting lower level constituents until individual preterminal symbols are written | b) Begins by hypothesizing a sentence (the symbol S) and successively predicting upper level constituents until individual preterminal symbols are written |
| c) Begins by hypothesizing lower level constituents and successively predicting a sentence (the symbol S) | d) Begins by hypothesizing upper level constituents and successively predicting a sentence (the symbol S) |
- (ii) Which of the following can be applied to input data sets based only on information in the training set?
- | | |
|----------------|-------------------------|
| a) postProcess | b) preProcess |
| c) process | d) All of the Mentioned |
- (iii) Select the the factors which affect the performance of learner system does not include.
- | | |
|-------------------------------|-------------------------|
| a) Representation scheme used | b) Training scenario |
| c) Type of feedback | d) Good data structures |
- (iv) In language understanding, identify the levels of knowledge that does not include
- | | |
|-----------------|--------------|
| a) Phonological | b) Syntactic |
| c) Empirical | d) Logical |
- (v) identify a model of language consists of the categories which does not include
- | | |
|-----------------------|----------------------------|
| a) Language units | b) Role structure of units |
| c) System constraints | d) Structural units |
- (vi) The action 'STACK(A, B)' of a robot arm specify for applying to _____
- | | |
|---|---|
| a) . Place block B on Block A | b) Place blocks A, B on the table in that order |
| c) Place blocks B, A on the table in that order | d) Place block A on block B |
- (vii) identify the library which is used for boosting generalized additive models?
- | | |
|-------------|-------------------------|
| a) gamBoost | b) gbm |
| c) ada | d) All of the Mentioned |

- (viii) identify characteristic of best machine learning method
- | | |
|-------------|-------------------------|
| a) Fast | b) Accuracy |
| c) Scalable | d) All of the Mentioned |
- (ix) Data used to optimize the parameter settings of a supervised learner model can be stated
- | | |
|---------------|-----------------|
| a) training | b) testing |
| c) validation | d) verification |
- (x) In Machine learning, dataset is divided into
- | | |
|---------------------------|-----------------------------|
| a) Training and Testing | b) Training and Forecasting |
| c) Training and Analyzing | d) None of these |
- (xi) $Y=mx+c$. Here m is described as
- | | |
|------------------|------------------|
| a) Y intercept | b) X intercept |
| c) Slope of line | d) None of these |
- (xii) identify the data which is used to build a data mining model
- | | |
|--------------------|------------------|
| a) Validation data | b) Testing Data |
| c) Training data | d) None of these |
- (xiii) Suppose you want to project high dimensional data into lower dimensions. The two most famous dimensionality reduction algorithms used here are PCA and t-SNE. Let's say you have applied both algorithms respectively on data "X" and you got the datasets "X_projected_PCA", "X_projected_tSNE". Select the true for "X_projected_PCA" & "X_projected_tSNE"
- | | |
|---|--|
| a) X_projected_PCA will have interpretation in the nearest neighbour space. | b) X_projected_tSNE will have interpretation in the nearest neighbour space. |
| c) Both will have interpretation in the nearest neighbour space | d) None of them will have interpretation in the nearest neighbour space. |
- (xiv) Suppose, you are given three variables X, Y and Z. The Pearson correlation coefficients for (X, Y), (Y, Z) and (X, Z) are C1, C2 & C3 respectively. Now, you have added 2 in all values of X (i.e. new values become X+2), subtracted 2 from all values of Y (i.e. new values are Y-2) and Z remains the same. The new coefficients for (X,Y), (Y,Z) and (X,Z) are given by D1, D2 & D3 respectively. establishe the relation between of D1, D2 & D3 and C1, C2 & C3
- | | |
|--------------------------------|--------------------------------|
| a) $D1= C1, D2 < C2, D3 > C3$ | b) $D1 = C1, D2 > C2, D3 > C3$ |
| c) $D1 = C1, D2 = C2, D3 = C3$ | d) $D1 = C1, D2 > C2, D3 < C3$ |
- (xv) predic the regression tasks
- | | |
|--|---|
| a) Predict the age of a person | b) Predict the country from where the person comes from |
| c) Predict whether the price of petroleum will increase tomorrow | d) Predict whether a document is related to science |

Group-B

(Short Answer Type Questions)

3 x 5=15

2. Explain a popular dimensionality reduction algorithm. (3)
 3. Explain Bias and variance with a suitable example and their impact. (3)
 4. define the issues / weekness of decision tree learning (3)
 5. Explain Back propagation algorithm. (3)
 6. Explain Variant and its effect. (3)
- OR**
- Expalin Overfit and its effect. (3)

Group-C

(Long Answer Type Questions)

5 x 6=30

- 7. Determine some differences between Random Forest and Gradient Boosting machines. (5)
- 8. Estimate the information Gain in case of decision tree with a suitable example. (5)
- 9. Define life cycle of machine learning. (5)
- 10. Illustrate stochastic Gradient Descent with example (5)
- 11. Explain the derivation of K-means algorithm. When it is recommended? (5)
- 12. Choose the best option - Random Forest and Naive-Baise Classifier. (5)

OR

Distinguish between Gradient Descent and Batch Gradient Descent. (5)

LIBRARY
Brainware University
Barasat, Kolkata -700125