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)

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) Like the probabilistic view, the _____ view allows us to associate a probability of membership with each classification

   a) Exemplar   b) Deductive
   c) Classical   d) Inductive

(ii) Select true statement about Naive Bayes?

   a) Assumes that all the features in a dataset are equally important
   b) Assumes that all the features in a dataset are independent
   c) Both Assumes that all the features in a dataset are equally important and Assumes that all the features in a dataset are independent
   d) None of these

(iii) A sentence parser typically is applied for

   a) It is used to parse sentences to check if they are utf-8 compliant.
   b) It is used to parse sentences to derive their most likely syntax tree structures
   c) It is used to parse sentences to assign POS tags to all tokens.
   d) It is used to check if sentences can be parsed into meaningful tokens.

(iv) Let us say that we have computed the gradient of our cost function and stored it in a vector g. Select the cost of one gradient descent updategiven the gradient?

   a) O(D)   b) O(N)
   c) O(ND)   d) O(ND2)

(v) select the widely used and effective machine learning algorithm based on the idea of bagging

   a) Decision Tree   b) Regression
   c) Classification   d) Random Forest

(vi) Machine learning is applied to

   a) Plot data   b) Analysis data
   c) Finding pattern in data   d) None of these

(vii) Supervised machine learning is divided into two groups

    a) Supervised and Unsupervised ML   b) Supervised and Nonsupervised ML
    c) Classification and Regression   d) None of these

(viii) Support Vector machine is observed as a

    a) Clustering algorithm   b) Feature Selection Algorithm
    c) Classification algorithm   d) None of these

(ix) K-Means is observed as a

    a) Clustering algorithm   b) Feature Selection Algorithm
    c) Classification algorithm   d) None of these

(x) Define Machine learning?

    a) The autonomous acquisition of knowledge   b) The autonomous acquisition of knowledge
       through the use of computer programs   through the use of manual programs
    c) The selective acquisition of knowledge   d) The selective acquisition of knowledge
       through the use of computer programs   through the use of manual programs

(xi) Select the correct one based on "Type-1" and "Type-2" errors? 1. Type1 is known as false positive and Type2 is known as false negative. 2. Type1 is known as false negative and Type2 is known as false positive. 3. Type1 error occurs when we reject a null hypothesis when it is actually tr

    a) . Only 1   b) Only 2
    c) Only 3   d) 1 and 3

(xii) In ensemble learning, you aggregate the predictions for weak learners, so that an ensemble of these models will give a better prediction than prediction of individual models. Select the true on the basis of the fact, weak learners used in ensemble model?1. They don't usually overfit.2. They have high bias, so they cannot solve complex learning problems3. They usually overfit.

    a) 1 and 2   b) 1 and 3
    c) 2 and 3   d) Only 1

(xiii) Select the good test dataset characteristic

    a) Large enough to yield meaningful results   b) Is representative of the dataset as a whole
    c) Both Large enough to yield meaningful   d)
       resultsand Is representative of the dataset   None of these
       as a whole

(xiv) A multiple regression model has

    a) Only one independent variable   b) More than one dependent variable
    c) More than one independent variable   d) None of these

(xv) A nearest neighborhood approach is best application

    a) for large size data set   b) When irrelevant attributes are removed from data
    c) When a generalized model of data is   d) When an explanation of what has been desirable   found is of primary importance

Group-B
(Short Answer Type Questions)
                                           3 x 5=15

2. Explain the limitations of Fixed Basis Function.   (3)
3. Explain the Naive Bayes classifier.   (3)
4. Describe the Bayesian Belief Network.   (3)
5. Describe confusion matrix. Explain it with an example.   (3)
6. Compare between perceptron and other ANN.   (3)
OR
Conclude the effectiveness of feed-back ANNs.   (3)

Group-C

7. Estimate the Gini Impurity in case of decession tree with a suitable example.   (5)
8. Justify that Kernel trick is making SVM a more powerfull method than linear regression. (5)
9. Define precision, recall and F1 Score   (5)
10. Explain the multiple linear regression with a suitable example.   (5)
11. Explain the term instance-based learning. Explain Soft Margin Classification   (5)
12. Develop the cost function of univariate linear regression and justify its meaning.   (5)

OR

justify that linear regression is a special case of polynomial regression with degree 1   (5)

****