BRAINWARE UNIVERSITY

Term End Examination 2023
Programme – M.Tech.(CSE)-AIML-2022
Course Name – Machine Learning
Course Code - PCC-MCSM201
(Semester II)

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)   Which of the following can be applied to inpute data sets based only on information in the training set?

a) postProcess
c) process

b) preProcess
d) All of the Mentioned

(ii) You run gradient descent for 15 iterations with a=0.3 and compute J(theta) after each iteration. You find that the value of J(Theta) decreases quickly and then levels off. Based on this, select the most justified statement

a) Rather than using the current value of a, use a larger value of a (say a=1.0)
c) a=0.3 is an effective choice of learning rate

b) Rather than using the current value of a, use a smaller value of a (say a=0.1)
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.
c) It is used to parse sentences to assign POS tags to all tokens.

b) It is used to parse sentences to derive their most likely syntax tree structures
d) It is used to check if sentences can be parsed into meaningful tokens.

(iv) Suppose you have trained a logistic regression classifier and it outputs a new example x with a prediction $h_\theta(x)$ = 0.2. This measure

a) Our estimate for P(y=1 | x)
c) Our estimate for P(y=1 | x)

b) Our estimate for P(y=0 | x)
d) Our estimate for P(y=0 | x)

(v)   Evaluate rank of the following matrix A= 1 1 1 1 1 1 1 1 1

a) 2
c) 3

b) 1
d) 4

(vi) Grid search is,

a) Linear in D.
b) Polynomial in D.
c) Exponential in N.
d) Linear in N.

(vii) Select the advantage of Grid search

a) It can be applied to non-di_erentiable functions.
b) It can be applied to non-continuous functions.
c) It is easy to implement.
d) All of these

(viii) Describe the cost parameter in the SVM

a) The number of cross-validations to be made
b) The kernel to be used
c) The tradeoff between misclassification and simplicity of the model
d) None of these

(ix) Anticipate the situation of happening when you use very small C (C~0)?

a) Misclassification would happen
b) Data will be correctly classified
c) Can't say
d) None of these

(x) We usually use feature normalization before using the Gaussian kernel in SVM. Select the most justified statement about feature normalization. 1. We do feature normalization so that new feature will dominate other 2. Sometimes, feature normalization is not feasible in case of categorical variables 3. Feature normalization always helps when we use Gaussian kernel in SVM

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

(xi) A feed-forward neural network is said to be fully connected when

a) mean squared error
b) root mean squared error
c) mean absolute error
d) mean relative error

(xii) 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

(xiii) identify the library which is used for boosting generalized additive models?

a) gamBoost
b) gbm
c) ada
d) All of the Mentioned

(xiv) 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

(xv) Data points have positive residual. Please selecte correct option

a) if they are above the regression line
b) if they are below the regression line
c) if the regression line actually passes through the point
d) None of the above

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

2. Explain Back propagation algorithm.   (3)
3. Interpret the ensemble models? Explain how ensemble techniques yield better learning as   (3)
   compared to traditional classification ML algorithms?
4. Discuss the differences between Lasso and Ridge regression?   (3)
5. Propose the concept of an ANN with a neat diagram.   (3)
6. Justify that Dimentionality reduction techniques like PCA reduces the training time and   (3)
   improves accuracy of a machine learning model.

OR

Discuss the major drawbacks of K-nearest Neighbour learning Algorithm and how it can be corrected   (3)

Group-C
(Long Answer Type Questions)   5 x 6=30

7. Define precision, recall and F1 Score   (5)
8. Illustrate Time series pattern and Horizontal Pattern. What are the goals of time series   (5)
    analysis and explain
9. Solve : the Cluster with following eight points (with (x, y) representing locations) into three   (5)
    clusters: A1(2, 10), A2(2, 5), A3(8, 4), A4(5, 8), A5(7, 5), A6(6, 4), A7(1, 2), A8(4, 9). Initial
    cluster centers are: A1(2, 10), A4(5, 8) and A7(1, 2).
10. Conclude the k-Means Algorithm with an example.   (5)
11. Conclude the concept of Bayes theorem with an example.   (5)
12. Conclude that SVM is better in some situation than Logistic regression.   (5)
    OR
    Analuyze and Compare between Logistic Regression and SVM.   (5)

****