SRM UNIVERSITY
DEPARTMENT OF INFORMATION TECHNOLOGY
Cycle Test 2
IT-R1141 Machine Learning

Register Number:

Degree : B.Tech
Year/Sem: 4/7
Duration : 3 hrs

Specialisation: IT
Date: 25/10/2017
Max. Marks: 100

Instructional Objective(s) covered in this test:

IO1. Understanding a very broad collection of machine learning algorithms and problems.

IO2. To learn algorithmic topics of machine learning and mathematically deep enough to introduce the required theory.

Student Outcome:

i. An ability to use current techniques, skills, and tools necessary for computing practice.

i2) An ability to understand tools necessary for computing practice

l. An ability to effectively integrate IT-based solutions into the user environment.

l1) An ability in integrating previous knowledge and new information

l2) An ability to integrating information in to IT based solutions into user environment.

Mark Allotments:

Question No	Instructional Objective	Out come/Sub Outcome	Mark (s) Scored	Sub Outcome Total
1	IO2	i2		
2	IO2	i2		
3	IO2	i2		
4	IO2	i2		
5	IO2	i2		
6	IO1	l1		
7	IO1	l1		
8	IO1	l1		
9	IO1	l1		

SRM University
Department of IT
VII Semester
Cycle Test 2
IT-R1141 Machine Learning

Maximum Marks: 100
Time: 3 hours

Part A **20x1=20 marks**

(Answer All Questions)

Choose the Most appropriate Answer

1. Logistic Regression is Not
   A. Regression at all
   B. y, the response variable can take on only two values, 0 and 1
   C. using discrete values for y
   D. used for classification

2. Given that $h_0(x)= g(a+ b.x1 +cx2)$ and if a= -3, b=1, c=1, predict y=1 in the 2 dimensional plot of x1 and x2
   A. Above line x1+x2 >=3
   B. below line x1+x2 >=3
   C. Above line x1+x2 =3
   D. below line x1+x2 =3

3. The maximal margin hyperplane depends
   A. directly on the other observations
   B. directly on the support vectors
   C. directly on the support vectors and on the other observations
   D. directly on the support vectors or on the other observations

4. In estimating parameters using Gradient descent for Univariate Linear Regression
   A. Update Θ_0 before finding new Θ_1
   B. Update $\Theta_1$before finding new Θ_0
   C. There is no rule for updating Theta
   D. Update Θ_0 and Θ_1 simultaneously

5. While calculating the gradient descent we may get
   A. Positive slope always
   B. Negative slope always
   C. Neither positive nor negative slope
   D. Either positive or negative slope

6. K means clustering algorithm
   A. Randomly picks up K training examples to set K Centroids
   B. Does not require the number of clusters to be specified in advance
   C. Assigns cluster centroid only once
   D. Is not suitable for sentiment analysis

7. The value of K in K means clustering
   A. Is always 3
   B. Is always 4
   C. Elbow method could be used to find K
   D. Can be greater than the number of training examples

8. Hierarchical clustering algorithm
   A. Requires number of clusters to be specified in advance
   B. Fixes cluster centroid only once
   C. Uses dendogram to determine clusters
   D. Is not suitable for unsupervised learning

9. Dendogram is NOT
   A. *bottom-up* or *agglomerative*
   B. Starting from trunk and combining clusters up to the leaves
   C. generally depicted as an upside-down tree
   D. Starting from the leaves and combining clusters up to the trunk

10. The motivation of dimensionality reduction is NOT
    A. Randomly ignoring some features
    B. Data compression
    C. Data Visualization
    D. Principal Component Analysis

11. Deep Feedforward Networks are NOT called
    A. Multilayer Perceptron
    B. Feedforward Neural Networks
    C. Artificial Neural Network
    D. Both A and B above

12. The f(x) of sigmoid function is given by
    A. 1/(1+exp(z))
    B. 1/(1-exp(z))
    C. 1/(1-exp(-z))
    D. 1/(1+exp(-z))

13. If $f(x)$ is linear, the ANN can draw
    A. Circular decision boundaries
    B. straight decision boundaries
    C. Non-linear decision boundaries
    D. No decision boundary

14. The use of hand-crafted features and relatively simple trainable is NOT
    A. Very tedious
    B. less expensive
    C. highly dependent on one application
    D. transferable easily to other applications

15. Therectified linear unit, or ReLU is defined by the activationfunction
A. g(z) = max{0, z}
B. g(z) = min{0, z}
C. g(z) = max{0, z^2}
D. g(z) = min{0, z^2}

16. A feature vector and corresponding parameter vector of an Artificial Neuron model was specified as follows:

$$X=\begin{pmatrix} x1 \\ x2 \\ x3 \end{pmatrix} \Theta= \begin{pmatrix} \theta 1 \\ \theta 2 \\ \theta 3 \end{pmatrix}$$

A. The specification is correct.
B. The specification is incorrect sinceΘ0 is missing.
C. The specification is incorrect sincex0 is missing.
D. The specification is incorrect sincex0 and Θ0 are missing.

17.An artificial neuron is NOT a device with
A. many inputs and many outputs.
B. training mode of operation
C. the using mode of operation
D. many inputs and one output.

18. The McCulloch and Pitts model (MCP)neuron fires if and only if;
A . X1W1 + X2W2 + X3W3 + ... = preset Threshold value T
B. X1W1 + X2W2 + X3W3 + ... <preset Threshold value T
C. X1W1 + X2W2 + X3W3 + ... > preset Threshold value T
D. X1W1 + X2W2 + X3W3 + ... == preset Threshold value T

19. Feed-forward NNs
A. allow signals to travel in both directions
B. allow signals to travel one way only
C. have one feedback loop
D. more than one feedback loop

20. Feedback networks are NOT
A. Dynamic; their 'state' is changing continuously until they reach an equilibrium point.
B. referred to as interactive or recurrent
C. Feedforward Networks
D. having signals traveling in both directions by introducing loops in the network

Part B

Answer any Five Questions **5x4=20 marks**

21. What exactly is deep learning?
22. Explain what you understand by the term auto encoders.
23. Explain how the human Neuron works in the brain, with the help of a diagram.
24. What is Euclidean distance and its use in Machine Learning
25. Explain Multi-class classification with an example.
26. Explain Feature Scaling and Mean Normalization.
27. State Gradient Descent Algorithm for Multi-variate Linear Regression.

Part C

5x12=60

28. A. Describe the working of the three variants of Support Vector Machine.

OR

28. B. Describe Logistic Regression.

29. A. Explain the algorithm for K means Clustering

OR.

29. B. Describe the algorithm for Hierarchical Clustering.

30. A. The data points (x1, x2) of a two dimensional scatter plot is given below:

$$\binom{1}{2}\binom{2}{4}\binom{3}{2}\binom{4}{4}\binom{5}{8}$$

Find the first Principal Component.

OR

30. B. Explain Dimensionality Reduction and Linear Discriminant Analysis.

31. A. Draw the diagram of an Artificial Neural Network with three inputs x1, x2 1 and x3 with a hidden layer with three Neurons. Derive the equation for the output. Assume appropriate weights.

OR

31. B. Implement the following types of Non-Linear classification with ANN

XNOR, OR, AND, NOR

32. A. Describe Convolutional Neural Networks.

OR

32. B. Implement XOR Feed with deep Forward Network and explain the algorithm.