BIRLA INSTITUTE OF TECHNOLOGY, MESRA, RANCHI
(END SEMESTER EXAMINATION)

CLASS: MTECH/PRE-PHD **SEMESTER : I/NA**
BRANCH: CS/IT **SESSION : MO/18**

SUBJECT: CS506 MACHINE LEARNING

TIME: 3 HRS. **FULL MARKS: 60**

INSTRUCTIONS:
1. The question paper contains 5 questions each of 10 marks and total 50 marks.
2. Attempt all questions.
3. The missing data, if any, may be assumed suitably.
***4*. Before attempting the question paper, be sure that you have got the correct question paper.**
5. Tables/Data hand book/Graph paper etc. to be supplied to the candidates in the examination hall.

Q.1(a) (I) Define machine learning (ML). Briefly explain the types of learning. [2+2+2+1]
(II) What are the key tasks of ML?
(III) Explain the steps required for selecting the right ML algorithm.
(IV) Explain any one issue in ML.

Q.1(b) Explain the concepts behind Linear Regression technique. The following table shows the midterm and final exam grades obtained for students in a database course. [3]

Midterm Exam (X)	Final Exam (Y)
72	84
50	63
81	77
74	78
94	90
86	75
59	49
83	79
65	77
33	52
88	74
81	90

Use the method of least squares using regression to predict the final exam grade of a student who received 86 on the midterm exam.

Q.2(a) Suppose you train a logistic regression classifier and your hypothesis function H is: [4]
$h_\theta(x)=g(\theta_0+ \theta_1x_1+ \theta_2x_2)$ where, $\theta_0=6$, $\theta_1=0$, $\theta_2=-1$
Draw the figure which will represent the decision boundary as given by above classifier. If you replace coefficient of x_1 with x_2 what would be the output figure.

Q.2(b) What are the issues in decision tree induction? For the given data determine the entropy after classification using each attribute for classification separately and find which attribute is best as decision attribute for the root by finding information gain with respect to entropy of temperature as reference attribute. [6]

Sl. No.	Temperature	Wind	Humidity
1	HOT	WEAK	NORMAL
2	HOT	STRONG	HIGH
3	MILD	WEAK	NORMAL
4	MILD	STRONG	HIGH
5	COOL	WEAK	NORMAL
6	MILD	STRONG	NORMAL
7	MILD	WEAK	HIGH
8	HOT	STRONG	NORMAL
9	MILD	STRONG	NORMAL
10	COOL	STRONG	NORMAL

PTO

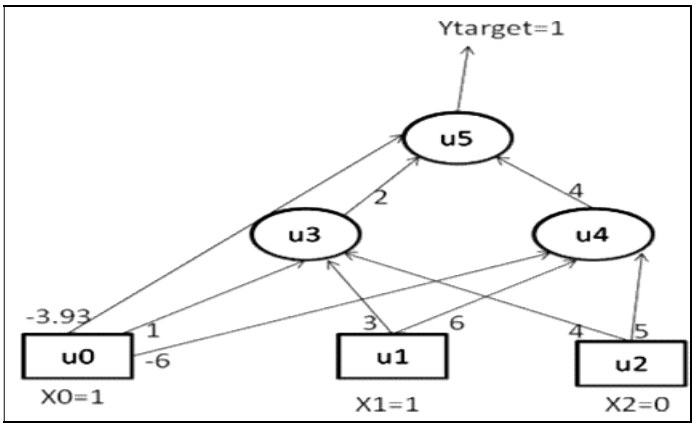
Q.3(a) What are the requirement of clustering algorithms? Apply k-means algorithm on given data set for K=2. Use $C_1(2,4)$ and $C_2(6,3)$ as an initial clusters. Data: a(2,4), b(3,3), c(5,5), d(6,3), e(4,3),f(6,6). [1+3]

Q.3(b) For the given set of points identify clusters using complete link and average link Agglomerative clustering algorithm and draw dendogram for each algorithm. [3+3]

	A	B
P1	1	1
P2	1.5	1.5
P3	5	5
P4	3	4
P5	4	4
P6	3	3.5

Q.4(a) (I) Develop a Perceptron for AND function with bi-polar inputs and targets. [1+2+2]
(II) Differentiate between Gradient Descent learning rule and Generalized Delta learning rule.
(III) What is the main advantage of using recurrent neural networks instead of feedforward neural networks? Explain the architecture of Elman and Jordon Simple Recurrent Neural Network.

Q.4(b) [4+1]



(I) Above figure is the example of Multilayer Perceptron NN (MLPNN). In this training example the input patterns are: $x_1=1$ and $x_2=0$ and output pattern is: $y_{target}=1$. Weight vectors are: $u_{13}=3$, $u_{14}=6$, $u_{23}=4$, $u_{24}=5$, $u_{35}=2$ and $u_{45}=4$. Bias weight: bias for unit 4 (u_4) $u_{04}=-6$. Similarly, bias weight $u_{03}=1$ and $u_{05}=-3.93$.
(a) Calculate the output and find out the error in first iteration.
(b) Update weight: with generalized delta learning rule evaluate weight vector after completion of one cycle of training (backpropagation).
(c) Again calculate the output for 2nd iteration and find out the error after 2nd iteration.
(d) Compare the new error with the previous error and check whether it has been reduced or not and above network is properly trained or not. Consider the learning constant is 0.1.
(II) Stop to train the neural network when overfitting is observed, i.e. when training data is being memorized -explain.

Q.5(a) Explain the followings: [2+2]
(I) Define support vector machine (SVM) and further explain the maximum margin linear separator concepts.
(II) Explain expected maximization (EM) for soft clustering.

Q.5(b) (I) What do you mean by BIAS and VARIANCE? [1+2+3]
(II) How ensembling helps in BIAS-VARIANCE trade off?
(III) Write short notes on Bagging.

:::::03/12/2018:::::M