Birla Institute of Technology & Science, Pilani
Work-Integrated Learning Programmes Division
First Semester 2018-2019
Comprehensive Examination (EC-3 Regular)

Course No.	: IS ZC464
Course Title	: MACHINE LEARNING
Nature of Exam	: Open Book
Weightage	: 50%
Duration	: 3 Hours
Date of Exam	: 25/11/2018 (FN)

No. of Pages = 2
No. of Questions = 10

Note:
1. Please follow all the *Instructions to Candidates* given on the cover page of the answer book.
2. All parts of a question should be answered consecutively. Each answer should start from a fresh page.
3. Assumptions made if any, should be stated clearly at the beginning of your answer.

Q.1. Explain the following concepts with respect to machine learning. [2 + 2 + 2 + 2 = 8]
(i) Self organizing map (SOM)
(ii) Lazy Learning verses Active Learning
(iii) Boosting
(iv) Perceptron and Neuron

Q.2. Device backpropagation rule for a neural network having **tanh** as activation function [6]

Q.3. Explain the concept of Vapnik-Chervonenkis (VC) dimension using shattering. How the number of training examples required to train the model is related to the VC dimension and what is its relation with training and test errors. [5]

Q.4. Naive Bayes classifier is a highly practical Bayesian learning method. It assumes that the attribute values are conditionally independent given the target value. Under this assumption, given a target value, the probability of observing the conjunction is just the product of the probabilities. For following given training set

Outlook	*Temperature*	*Humidity*	*Wind*	*PlayTennis*
Sunny	Hot	High	Weak	No
Sunny	Hot	High	Strong	No
Overcast	Hot	High	Weak	Yes
Rain	Mild	High	Weak	Yes
Rain	Cool	Normal	Weak	Yes
Rain	Cool	Normal	Strong	No
Overcast	Cool	Normal	Strong	Yes
Sunny	Mild	High	Weak	Yes
Sunny	Cool	Normal	Weak	Yes
Rain	Mild	Normal	Weak	Yes
Sunny	Mild	Normal	Strong	Yes
Overcast	Mild	High	Strong	No
Overcast	Hot	Normal	Weak	Yes
Rain	Mild	High	Strong	No

Determine classification of <Rain,Mild,High,Strong> explaining all important computation steps. [6]

Q.5. Sometime, when a child have a toothache it is due to cavity or gum problem. A full joint distribution of Toothache, cavity, and gum problem is given below.

	Toothache		No Toothache	
	Gum Problem	No Gum Problem	Gum Problem	No Gum Problem
Cavity	0.108	0.012	0.072	0.008
No Cavity	0.016	0.064	0.114	0.576

Define marginalization and determine a) the probability of cavity, b) probability of toothache in a child ? [3 + 3 = 6]

Q.6. Design a neuron for following classification function (specify weights). [5]

X1	X2	classification
0	0	1
0	1	1
1	0	0
1	1	0

Q.7. Apply uniform crossover on following chromosome with respect to the mask given. Show both the offspring in each case [4]

S.No.	Chromosome-1	Chromosome-2	Mask
1	10001010101	11100010001	00001110001
2	01010101010	10101010101	11001010010

Q.8. Recurrent neural networks (RNN) were an improvement over basic neural networks that can handle time dependent scenarios. Although it has been later discovered that they were not capable of learning long time dependencies. Long Short Term Memory (LSTM) were introduced to mitigate the issue of long term dependencies. Describe all the components of LSTM using a neat diagram and explain in detail how it is able to handle the issue of long term dependencies. [5]

Q.9. Describe how we can use genetic algorithms (GA) for optimization. What is its advantage? [2 + 2 = 3]

Q.10. Describe why it is important to visit all the states and take every action infinitely many times for the conversion in reinforcement learning. [2]

****