B.E.(CSE) Ass I ---CS6301 Machine Learning

Date: 24th NOV 2021 **Marks: 50** **Time: 1 ½ HR**

Answer All Questions

Part A (Answer Any 5) **5 X2 =10**

1. What is Machine Learning.? How is it different from Traditional Programming?
2. What is consistent hypothesis and version space?
3. What is bias and variance? Give its impact in choosing the best model.
4. Create a MLP network that solves the XOR function.
5. Compare LDA and PCA.
6. Consider (13, 14) is the location of a house and (25,20) is the location of the Market. Find the Euclidean distance and City-Block distance between the house and the Market.

Part B (Answer any 5) **5 x 8= 40**

1. Explain the use of machine learning for an application of your choice by specifying
   A. The Type of machine learning is needed to solve the problem.
   B. The Input and the expected output
   C. The Features to be used
   D. The possible machine learning algorithm to be used to solve the problem.
   E. The evaluation strategy to be used
2. Explain the various steps in the design of learning system and also discuss the design issues.
3. Compare and Contrast Supervised learning, Unsupervised learning, Semi supervised learning and Reinforcement learning. Give an example for each of learning.
4. Explain the concept behind RBF network and also with example explain how it is used for classification applications
5. Consider the following 2D data- (x,y):: (2.5 2.4) (0.5 0.7) (2.2 2.9) (1.9 2.2) (3.1 3.0) (2.3 2.7) (2 1.6) (1 1.1) (1.5 1.6) (1.1 0.9). Apply PCA or LDA to reduce the dimensions of the data.
6. Give the interpretation of Bayes rule. Find the class using naïve bayes classifier for the following data.

Example	Colour	Toughness	Fungus	Appearance	Poisonous
1	Green	Hard	N	Smooth	N
2	Green	Hard	Y	Smooth	N
3	Brown	Soft	N	Wrinkled	N
4	Orange	Hard	N	Wrinkled	Y
5	Green	Soft	Y	Smooth	Y
6	Green	Hard	Y	Wrinkled	Y
7	Orange	Hard	N	Wrinkled	Y
8	Green	Soft	Y	Wrinkled	?

CS6301 MACHINE LEARNING

3.12.2021 **Tutorial-1** **Marks : 10** **Duration : 50 mins**

1. Consider this training data set. Apply k-Means Clustering to this data set for k=2, i.e., you will produce two data clusters. Suppose you are given initial assignment cluster center as {cluster1: #1}, {cluster2: #10} – the first data point is used as the first cluster center and the 10-th as the second cluster center. Simulate the k-means (k=2) algorithm for ONE iteration. What cluster assignments after ONE iteration? Assume k-means uses Euclidean distance. What are the cluster assignments until convergence?

Data #	x	y
1	1.90	0.97
2	1.76	0.84
3	2.32	1.63
4	2.31	2.09
5	1.14	2.11
6	5.02	3.02
7	5.74	3.84
8	2.25	3.47
9	4.71	3.60
10	3.17	4.96

2. Consider the dataset below to learn a decision tree which predicts if people pass machine learning (Yes or No), based on their previous GPA (High, Medium, or Low) and whether or not they studied.
   a. What is the entropy H(Passed)?
   b. What is the entropy H(Passed / GPA)?
   c. What is the entropy H(Passed / Studied)?
   d. Draw the full decision tree that would be learned for this dataset.

GPA	Studied	Passed
L	F	F
L	T	T
M	F	F
M	T	T
H	F	T
H	T	T