Duration 3 Hours **[Maximum Marks 80]**

NOTE: -1) Question 1 is compulsory
2) Solve any three from the remaining five questions
3) Assume suitable data if necessary.
4) Figures to the right indicate full marks

Q.1. Attempt any four **20**

a) Explain loss function in machine learning.
b) Define entropy. Explain the use of entropy while forming a decision tree?
c) Explain dimensionality reduction techniques.
d) Specify functional difference between supervised and unsupervised learning.
e) Justify there is dilemma in Bias variance tradeoff for model selection.

Q.2.a) The pairwise distance between 6 points is given below. Draw hierarchy of cluster created by single link clustering algorithm. **10**

	P1	P2	P3	P4	P5	P6
P1	0	3	11	12	5	4
P2	3	0	9	8	10	9
P3	11	9	0	1	6	7
P4	12	8	1	0	7	8
P5	5	10	6	7	0	2
P6	4	9	7	8	2	0

b) Discuss in detail steps in K means clustering with one example. **10**

Q.3 a) Explain how machine learning is used in retail store sales prediction. **10**
b) Describe Support Vector Machine algorithm. **10**

Q.4 a) Explain the steps used in forming Classification and Regression trees. **10**
b) Given below is the confusion matrix for multiclass classification.
Find Precision, Recall and F1 Score for each class **10**

		True Class		
		Apple	**Orange**	**Mango**
Predicted Class	**Apple**	7	8	9
	Orange	1	2	3
	Mango	3	2	1

Q.5. a) List steps involved in Decision tree. State importance of pruning. **10**

b) Explain Baye's theorem. **10**

You have a burglar alarm installed at home. It is fairly reliable at detecting burglary but also sometimes responds to minor earthquakes. There are two neighbors, Tom and Jim, who promised to call you at work when they hear the alarm. Tom likes loud music and sometimes misses the alarm. Jim always calls when he hears the alarm but sometimes confuses telephone ringing with the alarms and calls too. Given evidence who has or has not called find probability that alarm has sounded but neither a burglary nor an earthquake has occurred and both Tom and Jim called.

P(Burglary) = 0.002, P(Earthquake) = 0.003,

B	E	P(A|B,E)
T	T	0.95
T	F	0.94
F	T	0.29
F	F	0.002

A	P(T|A)
T	0.8
F	0.01

A	P(J|A)
T	0.9
F	0.05

Q.6. a) Explain the concept of under fitting, over fitting and perfect fitting with suitable diagrams. How to avoid under fitting and overfitting **10**

b) Explain different error measures used for performance of regression.

Find Mean absolute and root mean square error for the following dataset. **10**

NO OF HOURSPENT DAILY FOR STUDY	MORKS OBTAINED	MARKS PREDICTED
1	61	60
1.5	65	62
2	71	70
2.5	75	80
3	85	90