

SECTION A

(20 Marks)

All questions are compulsory in this Section. All Questions are of 4 Marks.

Q.N.	Questions	Marks	CO	BL																																																							
1	What insights and competitive advantages can be obtained by adding Machine Learning into Finance sector?	4	CO1	L1																																																							
2	How would you evaluate a machine learning model? Discuss different metrics to evaluate the model.	4	CO2	L2																																																							
3	Explain how a decision tree splits nodes and the criteria it uses for splitting.	4	CO3	L2																																																							
4	Using Naïve Baye's Classification technique, classify {Red, Domestic, SUV} from the following car theft data: <table border="1" data-bbox="243 884 776 1183"> <thead> <tr> <th>Example No.</th><th>Color</th><th>Type</th><th>Origin</th><th>Stolen?</th></tr> </thead> <tbody> <tr><td>1</td><td>Red</td><td>Sports</td><td>Domestic</td><td>Yes</td></tr> <tr><td>2</td><td>Red</td><td>Sports</td><td>Domestic</td><td>No</td></tr> <tr><td>3</td><td>Red</td><td>Sports</td><td>Domestic</td><td>Yes</td></tr> <tr><td>4</td><td>Yellow</td><td>Sports</td><td>Domestic</td><td>No</td></tr> <tr><td>5</td><td>Yellow</td><td>Sports</td><td>Imported</td><td>Yes</td></tr> <tr><td>6</td><td>Yellow</td><td>SUV</td><td>Imported</td><td>No</td></tr> <tr><td>7</td><td>Yellow</td><td>SUV</td><td>Imported</td><td>Yes</td></tr> <tr><td>8</td><td>Yellow</td><td>SUV</td><td>Domestic</td><td>No</td></tr> <tr><td>9</td><td>Red</td><td>SUV</td><td>Imported</td><td>No</td></tr> <tr><td>10</td><td>Red</td><td>Sports</td><td>Imported</td><td>Yes</td></tr> </tbody> </table>	Example No.	Color	Type	Origin	Stolen?	1	Red	Sports	Domestic	Yes	2	Red	Sports	Domestic	No	3	Red	Sports	Domestic	Yes	4	Yellow	Sports	Domestic	No	5	Yellow	Sports	Imported	Yes	6	Yellow	SUV	Imported	No	7	Yellow	SUV	Imported	Yes	8	Yellow	SUV	Domestic	No	9	Red	SUV	Imported	No	10	Red	Sports	Imported	Yes	4	CO4	L4
Example No.	Color	Type	Origin	Stolen?																																																							
1	Red	Sports	Domestic	Yes																																																							
2	Red	Sports	Domestic	No																																																							
3	Red	Sports	Domestic	Yes																																																							
4	Yellow	Sports	Domestic	No																																																							
5	Yellow	Sports	Imported	Yes																																																							
6	Yellow	SUV	Imported	No																																																							
7	Yellow	SUV	Imported	Yes																																																							
8	Yellow	SUV	Domestic	No																																																							
9	Red	SUV	Imported	No																																																							
10	Red	Sports	Imported	Yes																																																							
5	You are tasked with predicting stock prices. Which machine learning algorithms would you consider, and why?	4	CO5	L3																																																							

SECTION B

(40 marks)

Attempt ANY FOUR Questions from this Section.

All Questions are of 10 Marks.

All Questions are of 10 Marks.																	
Q.N.	Questions		Marks														
6	For the given Table Transaction dataset, generate rules using Apriori Algorithm. Consider the values for Support=3 and Confidence = 75 %	<table><tr><th>Transaction ID</th><th>Itemset</th></tr><tr><td>T1</td><td>M, N, K, O, E, Y</td></tr><tr><td>T2</td><td>D, N, K, O, E, Y</td></tr><tr><td>T3</td><td>M, K, A, E</td></tr><tr><td>T4</td><td>M, C, K, U, Y</td></tr><tr><td>T5</td><td>C, O, I, K, E</td></tr></table>	Transaction ID	Itemset	T1	M, N, K, O, E, Y	T2	D, N, K, O, E, Y	T3	M, K, A, E	T4	M, C, K, U, Y	T5	C, O, I, K, E	10	CO1	L4
Transaction ID	Itemset																
T1	M, N, K, O, E, Y																
T2	D, N, K, O, E, Y																
T3	M, K, A, E																
T4	M, C, K, U, Y																
T5	C, O, I, K, E																
7	(a) Describe the differences between classification and regression tasks in machine learning, along with suitable examples. (b) Define Support factor and Confidence factor with reference to Association rule mining.		10	CO2	L2												

8	Construct a Decision Tree and write the classification rule for the following weather data:	10	CO3	L4																																																																											
	<table><tr><th>Outlook</th><th>Temperature</th><th>Humidity</th><th>Windy</th><th>Play</th></tr><tr><td>Sunny</td><td>Hot</td><td>High</td><td>False</td><td>No</td></tr><tr><td>Sunny</td><td>Hot</td><td>High</td><td>True</td><td>No</td></tr><tr><td>Overcast</td><td>Hot</td><td>High</td><td>False</td><td>Yes</td></tr><tr><td>Rainy</td><td>Mild</td><td>High</td><td>False</td><td>Yes</td></tr><tr><td>Rainy</td><td>Cool</td><td>Normal</td><td>False</td><td>Yes</td></tr><tr><td>Rainy</td><td>Cool</td><td>Normal</td><td>True</td><td>No</td></tr><tr><td>Overcast</td><td>Cool</td><td>Normal</td><td>True</td><td>Yes</td></tr><tr><td>Sunny</td><td>Mild</td><td>High</td><td>False</td><td>No</td></tr><tr><td>Sunny</td><td>Cool</td><td>Normal</td><td>False</td><td>Yes</td></tr><tr><td>Rainy</td><td>Mild</td><td>Normal</td><td>False</td><td>Yes</td></tr><tr><td>Sunny</td><td>Mild</td><td>Normal</td><td>True</td><td>Yes</td></tr><tr><td>Overcast</td><td>Mild</td><td>High</td><td>True</td><td>Yes</td></tr><tr><td>Overcast</td><td>Hot</td><td>Normal</td><td>False</td><td>Yes</td></tr><tr><td>Rainy</td><td>Mild</td><td>High</td><td>True</td><td>No</td></tr></table>	Outlook	Temperature	Humidity	Windy	Play	Sunny	Hot	High	False	No	Sunny	Hot	High	True	No	Overcast	Hot	High	False	Yes	Rainy	Mild	High	False	Yes	Rainy	Cool	Normal	False	Yes	Rainy	Cool	Normal	True	No	Overcast	Cool	Normal	True	Yes	Sunny	Mild	High	False	No	Sunny	Cool	Normal	False	Yes	Rainy	Mild	Normal	False	Yes	Sunny	Mild	Normal	True	Yes	Overcast	Mild	High	True	Yes	Overcast	Hot	Normal	False	Yes	Rainy	Mild	High	True	No			
Outlook	Temperature	Humidity	Windy	Play																																																																											
Sunny	Hot	High	False	No																																																																											
Sunny	Hot	High	True	No																																																																											
Overcast	Hot	High	False	Yes																																																																											
Rainy	Mild	High	False	Yes																																																																											
Rainy	Cool	Normal	False	Yes																																																																											
Rainy	Cool	Normal	True	No																																																																											
Overcast	Cool	Normal	True	Yes																																																																											
Sunny	Mild	High	False	No																																																																											
Sunny	Cool	Normal	False	Yes																																																																											
Rainy	Mild	Normal	False	Yes																																																																											
Sunny	Mild	Normal	True	Yes																																																																											
Overcast	Mild	High	True	Yes																																																																											
Overcast	Hot	Normal	False	Yes																																																																											
Rainy	Mild	High	True	No																																																																											
9	Apply k-means algorithm to make 3 clusters for given coordinates: (2,10), (2,6), (11,11), (6,9), (6,4), (1,2), (5,10), (4,9), (10,12), (7,5), (9,11), (4,6), (3,10), (3,8), (6,11)	10	CO4	L4																																																																											
10	(a) Discuss Cluster analysis. Also write the steps involved in determining k-means algorithm. (b) Predict the Class for Brightness = 20 and Saturation = 35, belongs to using KNN algorithm for below given dataset. <table><tr><td>Brightness</td><td>Saturation</td><td>Class</td></tr><tr><td>40</td><td>20</td><td>Red</td></tr><tr><td>50</td><td>50</td><td>Blue</td></tr><tr><td>60</td><td>90</td><td>Blue</td></tr><tr><td>10</td><td>25</td><td>Red</td></tr><tr><td>70</td><td>70</td><td>Blue</td></tr><tr><td>60</td><td>10</td><td>Red</td></tr><tr><td>25</td><td>80</td><td>Blue</td></tr></table>	Brightness	Saturation	Class	40	20	Red	50	50	Blue	60	90	Blue	10	25	Red	70	70	Blue	60	10	Red	25	80	Blue	10	CO5	L3																																																			
Brightness	Saturation	Class																																																																													
40	20	Red																																																																													
50	50	Blue																																																																													
60	90	Blue																																																																													
10	25	Red																																																																													
70	70	Blue																																																																													
60	10	Red																																																																													
25	80	Blue																																																																													
11	Write short notes on any TWO: (a) Artificial Neural Network (b) Machine Learning in Healthcare (c) (c) Supervised Learning (d) FP Growth Tree	10	CO1	L2																																																																											