Unit VI	
1)	Calculate the precision per class, recall per class, F1 score, accuracy, weighted average Precision and Weighted average Recall for the following dataset. | Tuple no.	1	2	3	4	5	6	7	8	9	10	11	12
---	---	---	---	---	---	---	---	---	---	---	---	---
Actual Label	P	P	P	P	N	N	P	N	N	P	P	N
Predicted Lebel	P	P	N	N	N	N	P	P	P	N	P	N	
2)	Discuss attribute selection criteria for the decision tree algorithm.
3)	Explain the support vector machine in detail. Apply SVM for XOR problem.
4)	How to improve the performance of the classifier model?
5)	Write a note on Ensembling techniques.
6)	Consider following three class confusion Matrix. Calculate Precision, Recall and F1-score for each class. Also calculate weighted average Precision and Weighted average Recall. Predicted 	Actual		Class 1	Class 2	Class 3
---	---	---	---	---
	Class 1	15	2	3
	Class 2	7	15	8
	Class 3	2	3	45	
7)	How to select the best value of K in KNN. Explain KNN steps.
8)	Consider the following 3-class confusion matrix. Solve the following Calculate the Weighted average Precision and Weighted Average Recall.

Predicted

Actual		A	B	C
	A	8	2	0
	B	1	9	0
	C	1	2	7

9) Give steps of Ensemble Learning. Explain majoring voting.

10) Construct Decision Tree upto three levels for following data.

Day	Weather	Temperature	Humidity	Wind	Play?
1	Sunny	Hot	High	Weak	No
2	Cloudy	Hot	High	Weak	Yes
3	Sunny	Mild	Normal	Strong	Yes
4	Cloudy	Mild	High	Strong	Yes
5	Rainy	Mild	High	Strong	No
6	Rainy	Cool	Normal	Strong	No
7	Rainy	Mild	High	Weak	Yes
8	Sunny	Hot	High	Strong	No
9	Cloudy	Hot	Normal	Weak	Yes
10	Rainy	Mild	High	Strong	No

11) Construct a decision Tree for the following data

Instance	A1	A2	A3	Classification
1	True	Hot	High	No
2	True	Hot	High	No
3	False	Hot	High	Yes
4	False	Cool	Normal	Yes
5	False	Cool	Normal	Yes
6	True	Cool	High	No
7	True	Hot	High	No
8	True	Hot	Normal	Yes
9	False	Cool	Normal	Yes
10	False	Cool	High	Yes

Apply KNN and find Label (Liked) for (Apple: false, Mango:True, Grapes: True) , Consider, K=3

	Apple	Mango	Grapes	Liked
A	True	True	True	False
B	True	False	False	True
C	False	True	True	False
D	False	True	False	True
E	True	False	False	True

13)

Apply Naïve Bais and find Label (Liked) for (Apple: false, Mango:True, Grapes: True)

	Apple	Mango	Grapes	Liked
A	True	True	True	False
B	True	False	False	True
C	False	True	True	False
D	False	True	False	True
E	True	False	False	True

14). Solve following Numerical example for Linear SVM:
Positively labelled data points (3,1)(3,-1)(6,1)(6,-1) and Negatively labelled data points (1,0)(0,1)(0,-1)(-1,0)

	Construct the Ensemble learning Model for following cases. Consider any example. By manipulating the training set, By manipulating the input features, By manipulating learning algorithms.
16)	Construct a decision tree using CART algorithm with the below training dataset shown in the table. | Instance	A1	A2	A3	Classification
---	---	---	---	---
1	True	Hot	High	No
2	True	Hot	High	No
3	False	Hot	High	Yes
4	False	Cool	Normal	Yes
5	False	Cool	Normal	Yes
6	True	Cool	High	No
7	True	Hot	High	No
8	True	Hot	Normal	Yes
9	False	Cool	Normal	Yes
10	False	Cool	High	Yes	

Unit V

1)	Explain the K-means algorithm in detail.
2)	Two clusters have the following data points. Calculate their intra cluster distance. Also calculate total inertia of the data point assignment. | Height	185	170	168	179	182	188	180	180	183	180	180	177
---	---	---	---	---	---	---	---	---	---	---	---	---
Weight	72	56	60	68	72	77	71	70	84	88	67	76
Assigned cluster	2	1	1	2	2	2	2	2	2	2	2	2	

	Differentiate between the Partitioning method and hierarchical method of clustering.
4)	Give advantages and disadvantages of Partitioned methods, Hierarchical methods, Density based method
5)	Consider flowing instances given as input to K-means clustering algorithm, k=3. Find members of these 3 clusters after two iterations X={(2,10), (2,5), (8,4), (5,8), (7,5), (6,4), (1,2), (4,9)}
6)	Compare Single link, Average link and Complete link clustering with example of any one type.
7)	Find the frequent itemsets and generate association rules on this. Assume that minimum support threshold (s = 2) and minimum confident threshold (c = 60%)
8)	Consider flowing instances given as input to K-means clustering algorithm, k=3. Find members of these 3 clusters after two iterations X={(2,10), (2,5), (8,4), (5,8), (7,5), (6,4), (1,2), (4,9)}
9)	Explain Market basket analysis using association rule and discussed the key terms associated with it.
10)	Transactional data is as follows:
11)	How Apriori algorithm work? Explain in detail.
12)	For the following set of transactions: (a1, a2, a3, a4)

For question 7):

Transaction ID	Items
T1	Hot Dogs, Buns, Ketchup
T2	Hot Dogs, Buns
T3	Hot Dogs, Coke, Chips
T4	Chips, Coke
T5	Chips, Ketchup
T6	Hot Dogs, Coke, Chips

For question 10):

TID	List of Item IDs
T1	I1, I2, I3
T2	I2, I4
T3	I2, I3
T4	I1, I2, I4
T5	I1, I3
Y6	I2, I3
T7	I1, I3
T8	I1, I2, I3, I5
T9	I1, I2, I3

Find the frequent itemset and generate the association rules with confidence values.

Or with FP growth algorithm.

	(a1, a4, a5, a6)	
	(a1, a2, a4, a7)	
	(a2, a3, a4, a7)	
	(a2, a3, a5, a6)	
	Find Association Rules for given support and confidence (3, 50%).	
13)	Give steps of the Apriori Algorithm	
14)	Compare Apriori and FP Growth.	
15)	What are the advantages of FP growth algorithm over Apriori?	
Unit VI		
1)	Apply Artificial neural network for disease prediction. Comment using performance metrics.	
2)	Give steps for building a Recommender System. Explain in detail.	
3)	Apply and analyse machine learning algorithm for medical diagnosis of diabetes based on patient data such as medical history, lab test results, and lifestyle factors. Comment on the type of algorithm can be applied and the performance evaluation parameters used for it.	
4)	Apply machine learning specifically for predicting the stock prices, analyze the performance evaluation parameters and the type of machine learning algorithm suitable for it.	
5)	How can machine learning be applied to Product Recommendation on E-commerce Websites? Explain steps involved in Product Recommendation using machine learning	
6)	Apply and analyse machine learning algorithm for medical diagnosis of diabetes based on patient data such as medical history, lab test results, and lifestyle factors. Comment on the type of algorithm can be applied and the performance evaluation parameters used for it.	
7)	Apply Artificial neural network for disease prediction. Comment using performance metrics.	
8)	Design a Recommender System for an ecommerce platform. Explain all steps involved in the process.	
9)	Apply machine learning specifically for predicting the stock prices, analyze the performance evaluation parameters and the type of machine learning algorithm suitable for it.	
10)	How can machine learning be applied to Product Recommendation on E-commerce Websites? Explain steps involved in Product Recommendation using machine learning	
11)	Which algorithm is best suited for stock prediction? Explain in detail.	
12)	Apply Artificial neural network for disease prediction. Comment using performance	

| | | metrics. |