Branch: CSE (III Year/ VI Semester)
6CS4-02-Machine Learning
Unit-2
TUTORIAL SHEET-1

Part-A: Short Answer Questions (Upto 25 words)

1. How would you define the number of clusters in a clustering algorithm?
2. What's the difference between probability and likelihood?
3. What is the main key difference between supervised and unsupervised machine learning?
4. There are many machine learning algorithms till now. If given a data set, how can one determine which algorithm to be used for that?
5. What is the significance of SSE in the context of k-means algorithm?
6. What are the demerits of Apriori Algorithm?
7. Define the following terms:
   a) Support count
   b) Item sets
   c) SSE

Part-B: Descriptive/ Analytical/ Design Questions

1.Explain in detail about Associative Rule Mining (ARM)?
2.Discuss various types of Unsupervised Machine Learning Techniques?
3.Discuss the strengths and weaknesses of the k-means algorithm and compare it with K-NN algorithm.
4.Differentiate between Apriori and FP growth algorithm with an example?
5.Discuss in detail the Gaussian Mixture Models.

Part-C Analytical/ Problem Solving Questions

1. With the help of Fuzzy-C means clustering, cluster the following points into two clusters 'C1' and 'C2':

S.No.	X	Y
1	5.1	3.5
2	4.9	3
3	4.7	3.2
4	4.6	3.1
5	5	3.6
6	5.2	2.7
7	6.6	2.9
8	4.9	2.4
9	6.3	3.3
10	5.7	2.8

2. Cluster the data points given in question-1, using the K-means algorithm.

3. For the transactions given in following table, illustrate the Apriori algorithm for finding frequent itemsets:

S.No.	Transaction number	List of items
1	T1	I1,I2,I5
2	T2	I2,I4
3	T3	I2,I3
4	T4	I1,I2,I4
5	T5	I1,I3
6	T6	I2,I3
7	T7	I1,I3
8	T8	I1,I2,I3,I5
9	T9	I1,I2,I3

4. For the item-transaction data table given in question-3, obtain the frequent itemsets using Frequent-growth pattern approach.