N.B. : 1) Question No.1 is **compulsory.**

     2) Attempt any **THREE** from the remaining questions.

     3) Figures to the right indicate full marks.

Q1.		Write a short note on :	
	(a)	Abstract Data types	[5]
	(b)	Regression Analysis	[5]
	(c)	Text Retrieval method	[5]
	(d)	Hierarchical clustering	[5]
Q2.	(a)	Explain parallel database architecture	[10]
	(b)	Explain Data Preprocessing in detail	[10]

Q3. (a) Differentiate following,   [10]

                  1. OLAP Vs OLTP

                  2. Star flake, snowflake and fact constellation Schema

     (b) Apply Apriori algorithm to the following data set to find out strong association   [10]
             rule with Support= 50% and Confidence=70%.

Transaction ID	Items
100	Fan, Tubelights, LED bulb , tape
200	Fan, Tubelights , LED bulb
300	Fan , Screws , capacitor
400	Fan , LED bulb , Screws
500	Tubelights , LED bulb , Screws

Q4.	(a)	Explain the KDD process in detail.	[10]
	(b)	Explain the Decision tree used in classification. Compare ID3, C4.5, CART classification algorithms	[10]
Q5.	(a)	Explain Data warehouse architecture in detail.	[10]

     (b) Generates the cluster using Euclidian distance for the given dataset using K-means clustering. (k=2). Consider X1 and X2 as seeds/centroids for two clusters respectively.

Item	Wt	Ht
X 1	1	2
X 2	1	1
X 3	2	3
X 4	3	4

                                                                  [10]

Q6. (a) Write a short note on   [10]

                  1. Bayes theorem

                  2. K-nearest neighbor classification

     (b) Explain the following,   [10]

                  1. Associative classification

                  2. Web mining