Q. P. Code: 40997

(3 HOURS) [TOTAL MARKS: ~~100~~ 80]

N.B. 1.Question 1 is compulsory
2.Attempt any ~~four~~ Three question out of the remaining six question.
3.All Question carry equal marks.
4.Illustrate answers with neat sketches whenever required.

Q.1(a) List and describe five primitives for specifying data mining task **10**
(b) Explain Data mining as a step in KDD. Give the architecture of typical Data Mining system. **10**

Q.2 (a) Explain BIRCH algorithm with example **10**
(b) Explain Hoeffding tree algorithm with example **10**

Q.3 (a) Explain Multilevel association rules with suitable example **10**
(b) Define classification , issues of classification and explain ID3 classification with example **10**

Q.4(a)Explain Data integration and data transformation w.r.t data warehouse **10**
(b) What is text mining ? Explain different approaches to text mining **10**

Q.5 (a) Explain Buisness Intelligence Issues **10**
(b) What is clustering? Explain k- means clustering algorithm. Suppose the data for clustering - {2,4,10,12,3,20,11,25}
Consider k=2, cluster the given data using above algorithm. **10**

Q.6(a) Explain sequence mining in Transactional databases **10**
(b) Explain periodic crawler and incremental crawler **10**
Q.7 Write short note on (Any two) **20**
(a) Web Usage mining
(b) Data Discretization and Summarization
(c) Spatial data cube and spatial OLAP

****