SUBJECT: (IT-704) Data Analysis & Information Extraction

Examination	: Block Sessional	**Seat No.**	: _____
Date	: 17/10/2014	**Day**	: Friday
Time	: 3:00 to 4:15	**Max. Marks**	: 36

INSTRUCTIONS:
1. Figures to the right indicate maximum marks for that question.
2. The symbols used carry their usual meanings.
3. Assume suitable data, if required & mention them clearly.
4. Draw neat sketches wherever necessary.

Q.1   Do as directed.

(a) What is a cluster analysis? A clustering in data-mining is an unsupervised learning. : **[2]**
    Justify

(b) Differentiate between query-driven approach and update-driven approach **[2]**

(c) What is bagging and boosting? Explain with example. **[2]**

(d) What do you mean by subject-oriented data in a data warehouse? **[2]**

(e) Define data mining. **[2]**

(f) Why do we require separate data warehouse? **[2]**

Q.2   Attempt *Any Two* from the following questions. **[12]**

(a) Explain Bayesian Classification Technique. **[6]**

(b) Explain BIRCH & CURE algorithm. **[6]**

(c) Explain any two partitions based clustering technique. **[6]**

Q.3   (a) Explain KDD process with figure. **[6]**

(b) Explain various types of concept hierarchy. **[6]**

OR

Q.3   (a) Differentiate between OLTP & OLAP. **[6]**

(b) Explain various technique of dimension reduction. **[6]**