

**SUBJECT CODE:8066**  
**FACULTY OF ENGINEERING AND TECHNOLOGY**  
**M.E.(Computer Science Engg.) Examination Nov/Dec 2015**  
**Data Mining & Big Data**  
**(Revised)**

[Time: Three Hours]

[Max. Marks: 80]

“Please check whether you have got the right question paper.”

N.B i) Solve any two questions from each section.

ii) Assume suitable data necessary and state it clearly.

**Section - A**

- Q1. a) A database has five transactions. Take min support count as 2. Find all frequent item sets using Apriori algorithm. 10
- |      |                    |
|------|--------------------|
| TID  | Items _ bought     |
| T100 | {M, O, N, K, E, Y} |
| T200 | {D, O, N, K, E, Y} |
| T300 | {C, A, K, E}       |
| T400 | {M, I, C, K, Y}    |
| T500 | {C, O, O, K, I, E} |
- b) What is constraint based association mining? What are the different type of constraints? 10
- Q.2 a) What is K-medoids algorithm? Explain with an example. How does it differ from k-Means algorithm? 10
- b) With an example explain how hierarchical clustering works using (i) Single linkage (ii) Complete linkage 10
- Q.3 a) What is the concept of page rank by which the page popularity is captured on web? Explain with the page rank algorithm. 10
- b) What is social network analysis (SNA)? How graph techniques are used for SNA? 10

**Section – B**

- Q.4 a) Which are the tools in Hadoop that can be used for machine learning and management & Deployment? 10
- b) A cloud uses 1500 nodes for data processing and has a processing of capacity 50 GB per hour. Consider the charges as 0.5 USD per node per hour, calculate the total cost and time required for processing one zettabyte of data on this cloud. 10
- Q.5 a) What is objective based data products? 10
- How will you apply drive train approach for marketing purpose?
- b) Explain “what it takes to build great machine learning products”? 10
- Q.6 a) Due to a huge collection of data in advance, is there any ‘dark side data’? Explain with an example. 10
- b) Describe in brief :- 10
- i) What to watch for in Big Data
  - ii) Temporal Mining