SUBJECT CODE NO:- P-8195
FACULTY OF ENGINEERING AND TECHNOLOGY
M.E. (Comp.Sci.& Engg.) Examination May/June 2017
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 question from each section
             ii) Assume suitable data, if necessary and state it clearly

Section A

Q.1   a) Describe Generation of Association rules from frequent item sets in Apriori method with an example   10
       database

       b) What is k-means algorithm? What are its limitations
                                                                                                                  10

Q.2   a) Describe agglomerative and divisive hierarchical clustering   10

       b) What are outliers? Describe any three methods of outlier detection. What is the importance of outliers? 10
Q.3   a) Explain the scale-free network with an example. What is a heavy-tailed distribution?   10

       b) What is Link Based Object Ranking (LBR)? What are the challenges in link mining?   10

Section B

Q.4   a) With an example describe the process of designing objective – based data product?   10

       b) Distinguish Clearly between Map Reduce and parallel DBMS Technology. Describe in detail an   10
       application of Big data technology
Q.5   a) Describe Apache Hadoop Architecture. Explain the role of coordination and workflow using zookeeper   10
       and Oozie

       b) How will you apply drive train approach for recommender system? Design and explain the steps   10
Q.6   a) What are the issues in mining Temporal sequences?   10

       b) Describe in brief: 10
       i) 'Personalization' and 'discrimination' in big data
       ii) Machine Learning products
       iii) web mining

2017