

## FACULTY OF ENGINEERING &amp; TECHNOLOGY

M.E (CSE)Examination - DEC - 2014

Data Mining and Big Data (Revised)

[Time: THREE Hours]

[Max. Marks: 80]

"Please check whether you have got the right question paper."

N.B

- 1) Solve any two questions from each section.
- 2) Assume suitable data, if necessary and state it clearly.

## SECTION A

- Q.1 A) How to improve efficiency of Apriori algorithm? 10  
 B) What is correlation analysis? The following contingency table summarizes transaction data. 10  
 Find the correlation between A and I using lift and cosine measures.

	A	B
I	400	350
$\bar{I}$	200	50

- Q.2 A) What are outliers? Describe any three methods of outlier detection. 10  
 B) The daily expenditures on Food (x1) and Travelling (x2) of five persons are shown below: 10  
 Give Data in format (Person, x1, x2)—(a, 2, 4), (b, 8, 2), (c, 9, 3), (d, 1, 5), (e, 8.5, 1)  
 i) Plot observations on a scatter diagram. How many clusters are formed visually?  
 ii) Calculate how the clusters will be formed using i) Single linkage ii) Complete linkage methods.  
 Use the Euclidean distance measure. Draw the dendrogram in each case with values.

- Q.3 A) What is spatial mining? Describe any one application which is based on it. 10  
 B) What is Page Rank Algorithm? How a web search takes place using Page Rank algorithm? 10  
 What are the characteristics of a social network?

## SECTION B

- Q.4 A) What are the characteristics of Big Data? What are its challenges? 10  
 B) Distinguish clearly between Map Reduce and Parallel DBMS technology. Explain with an example the costing associated with the cloud in processing big data. 10
- Q.5 A) Describe Apache Hadoop Architecture. What is the use Oozie, Zookeeper and Mahout tools? 10  
 B) How will you apply drive train approach for recommender system? Design and explain the steps. 10
- Q.6 A) Explain how 'Data Personalization' can be used for 'discrimination'. Justify with example. 10  
 B) Explain in brief how the application of 'spreadsheet' got enriched to 'dashboard'. What is the role of privacy in design? 10