FACULTY OF ENGINEERING

BE (CSE/IT)Final Year Examination - DEC – 2014

Data Warehousing & Data Mining(Revised)

[Time: THREE *Hours*]   [Max. Marks: 80]

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

N.B   1) Q.1 and Q.6 are compulsory solve any two from questions 2,3,4,5 and any two from questions 7, 8,9,10.
2) Assume suitable data if necessary and state it clearly.

SECTION A

Q.1	a)	What is the difference between OLTP and OLAP systems?	05
	b)	What are the various components in a multi-tiered Data warehouse architecture?	05
Q.2	a)	What are the important steps in data warehouse implementation?	08
	b)	Explain the structure of snowflake schema with a suitable database example. What is the use of data mining Query languages (DMQL)?	07
Q.3	a)	Define and describe any four OLAP operations on a sample data cube with examples and its result.	08
	b)	What are the major issues in data mining?	07
Q.4	a)	Define and describe the use of mean, mode, median, standard deviations in the representation of data object. What are the various distance measures used for dissimilarity between objects?	08
	b)	With a suitable block diagram, explain the principle of working of "Web Search Engine"	07

Q.5   Describe in brief (any three)   15
   a) Integration methods of DM system with DB/DW system.
   b) OLAP server architectures
   c) Data quality, data accuracy
   d) Tools for DW and DM implementation.

SECTION B

Q.6	a)	Define and describe classification and clustering of data.	05
	b)	What is market-basket analysis? What type of mining is possible from this data? What are the parameters used in the process?	05
Q.7	a)	Given the transactions below:	08

100{J,B}
200{M, Ch, Br, S}
300{C,Br}
400{ M, Br,S,P,A,T}
500 {Ch,S,Bf }
600{J,Br,K}

   i)   What is the support and confidence for following rules?
           Rule 1: M→Br   Rule 2: Br→M
   ii)   If the minimum support is set to 50% how many large item sets will be found?   07

   b) What is the process of generating association rules from frequent item sets? Describe with an example.

Q.8	a)	What are major applications of clustering? How do you cluster objects using K-meloids algorithm?	08
	b)	Explain in detail the phase of "Learning and Classification" in the working of any classifier.	07
Q.9	a)	Explain rule induction method using sequential covering algorithm. Explain the steps of rule generation in brief for a rule based classifier.	08
	b)	Explain the process of intelligence creation in BI. What is the importance of dashboards as a visualization tool?	07

Q.10   Describe in brief (any three)   15
   a)Decision tree classifier   b)Major tools used for BI
   c) Performance measurement of classifier   d) Methods for prediction.