INSTITUTE OF AERONAUTICAL ENGINEERING

(Autonomous)

Dundigal, Hyderabad - 500 043

Department of Computer Science and Engineering

ASSIGNMENT QUESTIONS

Course Title	DATA WAREHOUSING AND DATA MINING
Course Code	A70520 – JNTUH - R15
Class	IV B.Tech I Semester
Branch	CSE
Year	2018 – 2019
Course Faculty	Dr. K Suvarchala, Professor, CSE Dr. M. Madhubala, Professor, CSE Ms. B. Padmaja, Associate Professor, CSE Mr. P Anjaiah , Assistant Professor, CSE

OBJECTIVES

To meet the challenge of ensuring excellence in engineering education, the issue of quality needs to be addressed, debated and taken forward in a systematic manner. Accreditation is the principal means of quality assurance in higher education. The major emphasis of accreditation process is to measure the outcomes of the program that is being accredited.

In line with this, faculty of Institute of Aeronautical Engineering, Hyderabad has taken a lead in incorporating philosophy of outcome based education in the process of problem solving and career development. So, all students of the institute should understand the depth and approach of course to be taught through this question bank, which will enhance learner's learning process

S. No	QUESTION	Blooms Taxonomy Level	Course Outcome
	ASSIGNMENT – I **UNIT - I** **DATA WAREHOUSE**		
1	Explain with an example the different schemas for multidimensional databases?	Remember	1
2	Explain about the concept description? And what are the differences between concept description in large databases and OLAP?	Understand	2
3	Differentiate operational database systems and data warehousing?	Remember	2
4	Describe the three-tier data warehousing architecture?	Remember	2
5	Describe the complex aggregation at multiple granularities?	Remember	2
6	Discuss briefly about the data warehouse architecture?	Understand	1
7	Demonstrate the efficient processing of OLAP queries?	Understand	2
8	Compare the schemas for the multidimensional data models?	Understand	1
9	Explain the Data warehouse applications?	Understand	2
10	Discuss briefly about the multidimensional data models?	Remember	1
	UNIT – II **INTRODUCTION TO DATA MINING**		
1	Distinguish between the data warehouse and databases? How they are similar?	Understand	3
2	Describe three challenges to data mining regarding data mining methodology and user interaction issues?	Remember	3

3	Discuss briefly about the data smoothing techniques?	Understand	4
4	Explain Data Integration and Transformation?	Remember	4
5	Describe the various data reduction techniques?	Remember	4
6	Define data cleaning? Express the different techniques for handling missing values?	Understand	5
7	Explain data mining as a step in the process of knowledge discovery?	Understand	5
8	List and describe the five primitives for specifying a data mining task?	Remember	5
9	Explain the difference between discrimination and classification? Between characterization and clustering? Between classification and prediction? For each of these pairs of tasks, how are they similar?	Remember	5
10	Distinguish between the data warehouses and data mining?	Remember	4
UNIT – III **ASSOCIATION RULES**		
1	Define the terms frequent item sets, closed item sets and association rules?	Understand	5
2	Discuss which algorithm is an influential algorithm for mining frequent item sets for Boolean association rules? Explain with an example?	Remember	6
3	Describe the different techniques to improve the efficiency of	Remember	5
4	Discuss the FP-growth algorithm? Explain with an example?	Remember	6
5	Discuss about mining multilevel association rules from transaction databases in detail?	Remember	6
ASSIGNMENT – II		
6	Discuss about constraint-based association mining?	Remember	6
7	Discuss about mining multilevel association rules from transaction	Understand	6
8	Describe about the correlation analysis using Chi-square?	Remember	6
9	Explain what are additional rule constraints to guide mining?	Remember	6
10	Illustrate about the correlation analysis using All-certainty Measure?	Remember	6
UNIT – IV **CLASSIFICATION**		
1	Explain about the classification and prediction? Example with an example?	Understand	7
2	Discuss about basic decision tree induction algorithm?	Remember	7
3	Summarize how does tree pruning work? What are some enhancements to basic decision tree induction?	Understand	7
4	Explain how scalable is decision tree induction? Explain?	Remember	8
5	Describe the working procedures of simple Bayesian classifier?	Remember	8
6	Discuss the back propagation algorithm and explain?	Understand	8
7	Explain about classifier accuracy? Explain the process of measuring the accuracy of a classifier?	Understand	7
8	Explain training of Bayesian belief networks?	Remember	8
9	Explain briefly about the Navie Bayesian Classification?	Remember	7
10	Differentiate classification and prediction methods?	Remember	8
UNIT – V **CLUSTERING**		
1	Discuss the various types of data in cluster analysis?	Remember	10
2	Explain the categories of major clustering methods?	Remember	9
3	Write algorithms for k-means and k-medoids? Explain?	Remember	10
4	Describe the different types of hierarchical methods?	Remember	9
5	Discuss about the DBSCAN density-based methods?	Remember	10
6	Demonstrate about the following hierarchical methods?	Remember	10
7	Explain the working of CLIQUE algorithm?	Remember	9

8	Define the distance-based outlier? Illustrate the efficient algorithms for mining distance-based algorithm?	Remember	10
9	Explain about the Statistical-based outlier detection?	Understand	10
10	Explain about the agglomerative and divisive hierarchical methods?	Understand	9

Prepared By: Ms. K Suvarchala, Professor, CSE

HOD, CSE