Subject: Data Mining and Business Intelligence

1. Explain web mining using example.
2. Explain Hadoop architecture using figure.
3. Explain mapreduce. Explain any example using mapreduce.
4. Explain big data and big data analytics. Explain key roles and their responsibilities for successful analytic project.
5. Explain text mining using example.
6. How data Mining is useful for Business Intelligence applications viz.Balanced Scorecard, Fraud Detection, Clickstream Mining, Market Segmentation, retail industry, telecommunications industry, banking & finance and CRM
7. Explain Spatial mining using example.
8. Explain Mining Multiple-Level Association Rules using example.
9. Differentiate between Overfitting and Tree Pruning w.r.to following parameters. i). definition figure ii). use in particular situation 03 iii). Limitation
10. Explain KDD process using figure.
11. Do feature wise comparison between BI and DW.
12. Explain research issues in Data Mining.
13. Enlist the preprocessing steps with example. Explain procedure of any technique of preprocessing.
14. Do feature wise comparison between ROLAP and MOLAP.
15. Explain schemas: Stars, snowflakes and fact constellations using figures.
16. Explain various features of Data Warehouse?
17. Differentiate between Operational Database System and Data Warehouse 18. What is Concept Hierarchy? List and briefly explain types of Concept Hierarchy 19. Explain Mean, Median, Mode, Variance, Standard Deviation & five number summary with suitable database example.
20. What is noise? Explain data smoothing methods as noise removal technique to divide given data into bins of size 3 by bin partition (equal frequency), by bin means, by bin medians and by bin boundaries. Consider the data: 10, 2, 19, 18, 20, 18, 25, 28, 22 21. Define data cube and explain 3 operations on it.
22. Define following terms : Data Mart , Enterprise Warehouse & Virtual Warehouse
23. Draw the topology of a multilayer, feed-forward Neural Network.
24. Explain Linear regression with example.
25. Explain the major issues in data mining
26. Briefly explain text mining
27. What is market basket analysis? Explain the two measures of rule interestingness: support and confidence

28. What is Big Data? What is big data analytic? Explain the big data distributed file system.

29. Discuss various layers of multilayer feed-forward neural network with diagram.

30. What is apex cuboid? Discuss drill down and roll up operation with diagram. 31. Enlist data reduction strategies and explain any two.

32. Discuss attribute subset selection.

33. Can BI is used for DM? Or vice versa? Justify.

34. Draw and Explain Snowflakes and Fact constellations Schema.

35. Discuss Following: (i) Meta Data (ii) Virtual Warehouse

36. Describe Concept Hierarchy? List and briefly explain types of Concept Hierarchy

37. Define "clustering"? Mention any two applications of clustering. 38. Briefly explain Linear and Non-linear regression.

39. Apriori Algorithm. Example

40. In data pre-processing why we need data smoothing? Discuss data smoothing by Binning.