Set of Questions on Data Mining

1. What is Data Mining?
2. Identify the origins of Data Mining.
3. List the motivations for data mining.
4. List the tasks of data mining.
5. What are the applications of classification?
6. What are the applications of clustering?
7. What are the applications of outlier analysis?
8. What are the applications of association rule mining?
9. Define the following:
   1. Object
   2. Attribute
10. Describe the classification of attributes.
11. What is discrete and continuous attribute?
12. Enlist the important characteristics of data.
13. What is a market basket data? Explain.
14. Describe the sparse matrix?
15. Explain the graph based data.
16. Describe the Spatial Data
17. What is temporal coherence?
18. List the issues with data quality.
19. How do you handle the missing data?
20. Define a metric.
21. Describe the Pearson correlation coefficient.
22. What is aggregation? What the goal of aggregation in data mining?
23. Describe the sampling types.
24. What is binarization?
25. What is transformation? List the different transformation techniques.
26. What is the curse of dimensionality? Explain.
27. What are the design issues of decision tree induction?
28. Define: Gini Index
29. Define: Gain
30. List the different classifiers.
31. Explain the methods for expressing ordinal attribute test conditions.
32. Explain the methods for expressing nominal attribute test conditions.
33. Explain the methods for expressing continuous attribute test conditions.
34. What do you mean by descriptive modelling?
35. What do you mean by predictive modelling?
36. For the following data calculate the SMC: x={1,0,0,0,0,0,0,0,0,1}, y={0,0,0,0,0,0,1,0,1,0}
37. For the following data calculate the J: x={1,0,0,0,0,0,0,0,0,1}, y={0,0,0,0,0,0,1,0,1,0}
38. What is a document matrix? Explain.
39. Describe the general approach to classification task.
40. Explain the different methods for feature subset selection.