

Tutorial: Data Mining

1. What is Inductive learning? Examples are expected for supporting your discussion.
2. Discuss differences between classification and association rules methods.
3. Outline OneR algorithm. Critically discuss and evaluate its advantages and disadvantages, supported by examples.
4. Outline and explain ID3 algorithm. Critically discuss and evaluate its advantages and disadvantages, supported by examples.
5. Discuss the difference between ID3 and C4.5
6. Describe and discuss the two step implementation of the search for association rules.
7. Describe and discuss the Breadth first search (Apriori).
8. Frequent pattern mining was originally developed – in the special form of frequent item set mining. Use an example to explain the challenge of frequent pattern mining. Critically discuss and analyse possible solutions to this challenge.
9. Outline Apriori algorithm. Critically discuss and evaluate its advantages and disadvantages, supported by examples.
10. Analyse and assess key factors that are important for selecting rules from models generated by data mining tools.
11. Explain and discuss the two class problem.
12. Discuss and analyse the approach to this two class problem. Critically discuss the advantages and disadvantages of using regression to analyse two-class classification problems.
13. Briefly discuss the types of clustering approaches.
14. Discuss the approaches to choosing right clusters.