

Data Mining 2011

Examination June 9th, 2011

You may answer either in english or in finnish.

1. How would you lead a data mining project (in a company and application of your choice)?
2. What are *association rules*? What properties make an association rule a good one? What are the challenges in mining such rules from large datasets? Explain what kind of methods can be used to make association rule mining practical.
3. Explain the concepts of *bias* and *variance*, especially their relationship to various data mining methods. Discuss the possible ways of reducing variance in data mining tasks.
4. What are *bagging* and *boosting*? Explain these techniques and discuss their common properties, differences and (possible) performance gains.