Name

Machine Learning
Rijks*universiteit* Groningen

Fall
Instructor: Jörg Tiedemann

Exam
9:00 am Fri. 26 Oct. 2006
Academy Building, Zernike hall

Important Instructions

- Write your name & student number above, write initials on each page
- This is a three-hour exam.
- Answer all questions. There are 50 points in total and you will need at least 30 to pass the exam.

Questions

1. Name and briefly describe 3 applications where machine learning is useful! (3)
2. Describe how distance-weighted kNN can be used for classification with multiple (more than 2) discrete classes. (3)
3. Discuss the following argument: "K-fold cross validation uses the entire data set for training. It is therefore necessary to use a separate, untouched evaluation set for the final evaluation." (3)
4. What kind of inductive bias is used in kNN? Explain! (4)
5. What is the advantage of locally weighted regression over distance-weighted kNN for regression? (3)
6. If you would have to choose between a Naive Bayes classifier and kNN: Which approach would you select for learning the concept of playing tennis with features as described below? Explain why! (3)

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No

7. Decision tree learning implements ideas of the minimum description length principle. Explain why information gain is a useful measure to approach this principle. (4)

8. Is a kNN classifier a consistent learner? Why? Why not? (3)

9. Can instance based learning (k nearest neighbours) overfit? If yes, how can that be avoided? If no, why not? (3)

10. Does ID3 produce the smallest decision tree that is consistent with the training data? (explain) (3)

11. Explain how costs can be integrated into decision trees and why this is useful (give two examples)! (3)

12. What problem is addressed when using split information and how is it solved using this measure? (2)

13. Naive Bayes classifiers can be seen as Bayesian Belief Networks. Draw a graph for such a model and explain it using a practical example. (3)

14. What is a prior in Naive Bayes and what is it used for? (2)

15. How can real-valued data attributes be handled in Naive Bayes? (3)

16. What is a MAP hypothesis? (2)

17. You know that 1 out of 10 mails arriving at your mail-server is not a spam mail. Would you be impressed, if you hear that your spam filter has an accuracy of 85%? Why? Why not? Would your judgement change if you would know that the system yields a recall of 100% of detecting spam mails? (3)