
WRITTEN EXAMINATION
“Foundations of Machine Learning and Data Mining”
SUMMER TERM 2006
March 28TH, 2007
PROF. DR. RALF MÖLLER

Name, Student Id:

Name of curriculum:

Signature:

IT IS NOT PERMITTED TO WRITE ON THESE PIECES OF PAPER BEFORE THE START SIGNAL. IN ADDITION, WRITING ON THESE PIECES OF PAPER AFTER THE END SIGNAL WILL HAVE THE EFFECT THAT YOU DO NOT PASS THIS EXAM WITHOUT FURTHER WARNINGS. THIS ALSO APPLIES IF YOU WRITE YOUR NAME AND/OR STUDENT ID AFTER THE END SIGNAL.

Please note:

- 1. Do not open the exam sheets until permitted by the supervisor!**
- 2. Put your student identification card as well as your passport on the table.**
- 3. Check if your student Id is on the list of Ids handed out by the supervisors. If your Id is not on this list, then**
 - (a) Fill out a proviso, which will be given to you by the supervisor. You may not start the exam until you have returned the proviso to the supervisor.**
 - (b) Also fill out the additional form given to you by the supervisor. Please go to the students office which is responsible for your curriculum and let them sign the form. Personally return the signed form to the STS secretary (Harburger Schloßstr. 20, 2nd floor, Frau Hantschmann); bring your passport for identification as well.**
- 4. You have 90 minutes for answering the questions. Additional resources are not allowed.**
- 5. There is sufficient space for your solutions on the examination sheets.**
- 6. If you receive additional pieces of paper from the supervisor, please write your name and student Id also on these pages, and add a page number.**
- 7. If you need to leave the examination room, silently inform the supervisors. Do not leave your table. Wait until the supervisor approaches your desk and gives you the permission to leave the room.**

1) Describe in your own words why and where concepts of “Machine Learning” can be used in the context of “Enterprise Resource Planning”.

2) Show from the axioms of probability theory and the definition of conditional probability that $P(a|b \wedge a) = 1$.

3) The type of feedback available for learning is usually the most important factor in determining the nature of the learning problem the agent faces. The field of machine learning usually distinguishes three cases: supervised, unsupervised and reinforcement learning. Describe these three terms and give an example for each case.

4) Suppose you have the following distribution $\mathbf{P}(A, B, C)$ with three Boolean variables.
a. How many independent entries does this distribution contain?

- b. Assume we have conditional independence of A and B given C. Transform the expression using the chain rule and the formula for conditional independence and determine the number of independent entries in the CPTs (Conditional Probability Tables).

- 5) In your car, there is an alarm that senses when a temperature gauge exceeds a given threshold. The gauge measures the temperature of the motor. Consider the Boolean variables A (alarm sounds), F_A (alarm is faulty), and F_G (gauge is faulty) and the multivalued nodes G (gauge reading) and T (actual motor temperature).

- a. Draw a Bayesian network for this domain, given that the Gauge is more likely to fail when the core temperature gets too high.

- b. Suppose there are just two possible actual and measured temperatures, normal and high; the probability that the gauge gives the correct temperature is x when it is working, but y when it is faulty. Give the conditional probability table associated with G and A .

- 6) A decision tree takes as input an object or situation described by a set of attributes and returns a “decision” – the predicted output value. Decision tree induction is one of the most successful forms of machine learning. To build a decision tree from a set of examples, one has to decide the structure of the decision tree.
 - a. The information gain of attributes can be used to decide which attribute should be next within the tree. Give the formula of the information gain and explain its meaning?

- b. Instead of the information gain, the gain ratio can be used to determine the next attribute in the decision tree. What kind of problem of the information gain is compensated by the gain ratio?

The gain ratio is given by: $\text{GainRatio}(D,T) = \text{Gain}(D,T) / \text{SplitInfo}(D,T)$

- 7) Explain the concept of Parameter Learning with complete data. For this task assume you have a Bayesian Network.

8) What is a version space and how is learning organized in this context. What is the general advantage of version space compared to decision trees?

9) In the context of relevance based learning.

a. Explain the term "determination" by giving an example.

b. Do determinations influence the size of the hypotheses space and in what direction?

10) In the context of “non-parametric learning” explain the basic idea of the “Nearest Neighbor Classifier”.

11) Does parametric learning cover the same amount of possible hypotheses as non-parametric learning?