

Information Theory and Machine Learning

T-61.182 Special Course in Computer and
Information Science II, Spring 2004 (4 cr)

Prof. Juha Karhunen
Helsinki University of Technology

General information

- The course is based on selected parts of the book

David J.C. MacKay: Information Theory, Inference, and Learning Algorithms.

- Publisher: Cambridge Univ. Press, 2003.
- About 640 pages \Rightarrow We shall skip parts of the book in this course.
- Price: Officially 30 UK pounds (about 45 euros).
- Prices in internet booksellers vary, in some cases lower than 30 UK pounds.
- Comments and reviews on the book have been quite positive.
- Understanding the book requires some mathematical maturity.

- We shall skip advanced special topics and non-essential theoretical proofs in our course.
- Home page of the book:
<http://www.inference.phy.cam.ac.uk/mackay/itila/>
- The book is freely viewable on the web site of the book in different formats (pdf, ps, djvu).
- There is also some other material like found errors, software, etc.
- The book deals with information theory, inference, and machine learning.
- These topics are usually discussed separately, but actually they are closely related.
- They are important in many areas of science and engineering:
- Communications, signal processing, data mining, pattern

recognition, learning theory, cryptography, and bioinformatics.

- In this course, we deal with basic information theory and coding (parts I and II).
- As well as Bayesian inference methods (part IV).
- Part V, neural networks, is largely overlapping with our two neural network courses.
- We deal briefly with things which open up fresh views or are not discussed in our neural network courses (if time allows).
- Advanced topics in information theory (Part III) and sparse graph coding (Part VI) are skipped in this special course.

Organization of the seminar

- Four (4) credit points as usual in our seminar courses.
- Weekly in the lecture room T4 on Thursdays 14:15 (up to 16 if necessary).
- A more detailed programme will be planned when the number of participants is known.
- Responsible teacher: Prof. **Juha Karhunen**
- Email: Juha.Karhunen@hut.fi, room TB327, tel. 451 3270.
- Course assistant: M.Sc. **Antti Honkela**
- His email: Antti.Honkela@hut.fi; room TB311.
- The language of the course is English due to foreign participants.
- The course is intended mainly for graduate students.

- Can be taken by undergraduate students who are mature enough.
- Sufficient mathematical background, most of M.Sc. studies done.
- You should know the fundamentals of probability theory and linear algebra.
- There will be no examination; would be difficult for graduate students working full-time in corporations etc.
- You should write your name, study book number, email address, and department to the enrollment list circulating.
- The price of the book is low, 30 GBP pounds officially.
- So it is recommended that you purchase the book.
- It takes some time until you will get the book.
- Please reserve your own talk from the list circulating in the seminar meetings as soon as possible.

Requirements for passing the course

- Sufficient participation (about 70%) in the seminar meetings.
 - Put a cross to the attendance list whenever you attend.
- You must prepare and present your own talk.
 - Usually one talk per meeting, about 45 min - 1 hour long.
 - If there are two talks, they could be about 35-45 min each.
 - Skip proofs, excessive theory and details, and try to explain clearly the most important matters of the scope of your talk.
 - You can give your slides (for example in .pdf format) to the course assistant.
 - He can copy them to the participants and/or put your slides onto the home page of the course.

- Solving sufficient amount of the given problems.
 - Solving 50% of the total number of problems suffices for the mark “accepted” (hyväksytty).
 - Solving 90-95% of problems is required for the mark “accepted with distinction (kiittäen hyväksytty)
 - Typically there will be 2 – 3 problems per each talk.
 - You should select yourself 2 – 3 problems on the portion of the book covered in your talk.
 - Some problems have been solved in the book; of course, you should not select them.
 - You must be able to solve these problems yourself!
 - This guarantees that the problems are not too difficult.
 - Give the correct solutions of the your problems to the course organizer Prof. Juha Karhunen (or to assistant Antti Honkela).

- Problems have been classified in the book: 1= simple, 2=medium, 3=moderately hard, 4=hard, 5=research project.
- You should select only problems having degrees 1, 2, or 3.
- Problems are useful because they force people to read the corresponding parts of the book.
- It is preferable but not necessary to return your solutions to the problems given within 2 weeks.
- Deadline for returning solved problems: May 15th, 2004.
- Somewhat open issue: replacing some problems by computer assignment(s) giving hands-on experience!?