T-61.6020: Machine Learning: Basic Principles (5 cr) V P

Kai Puolamäki

15.1.2007

http://www.cis.hut.fi/Opinnot/T-61.6020/

t616020@james.hut.fi

People

- Teacher (vastaava opettaja): Kai Puolamäki, PhD, lecturing researcher
- Assistant: Mikko Korpela, MSc
- You can contact us at the lectures (we also have offices in the third floor)
- Please use the course email address t616020@james.hut.fi in course-related correspondence

About this course

- Lecture/seminar in T4 every Monday 14-16 from 15 January to 30 April 2007 (no lectures in exam period on 5 March and Easter holiday on 9 April)
- If you pass, you get five credit points
- This is a post-graduate course (lisenssiaattikurssi)
- Language: English or Finnish
- Prerequisites:
  - First two years' mathematics courses
  - Suitable course for advanced undergraduates or PhD students
  - Some prior knowledge of the topics does not hurt

Requirements

- To pass the course you should:
  - Participate actively (over 70% presence at the lectures)
  - Prepare a presentation(s) on given topic(s)
  - Get at least 50% of the points from the problems
- To pass the course with distinction you should additionally:
  - Have a good presentation
  - Solve the exercise problems well (i.e., well over 50% of the points)

Contents

- The purpose of the seminar is to read through the new text book by Christopher M. Bishop, "Pattern Recognition and Machine Learning" (Springer 2006).
- This book will be used on a new yearly lecture course (with the same name as this seminar), beginning autumn 2007
- We will go one chapter per week:
  - Student(s) will prepare a presentation
  - Everyone will read the chapter before the presentation
  - All will solve some exercises before the presentation (they are submitted to the assistant at presentation)

Contents

Solving problems

- You should submit the solved problems at the lectures: only solutions submitted in time count in grading
- If you cannot make it to some lecture, you can submit the problems on previous lecture or directly to assistant
- The problems are announced (at least) two weeks before submission deadline
- We will (mainly) use the problems from Bishop's book
- The problems are classified with 1-3 stars (*="Simple exercise taking few minutes to complete", ***="Significantly more complex exercise")
- If you solved the problem correctly and well, you get one point for one star; a good attempt to correct direction will get half of the stars

Next problems

- 22 January. 10 stars/points from the following problems:
  - at least 2 stars from 1.1-1.41
  - at least 2 stars from 2.1-2.30
  - at least 2 stars from 2.31-2.60
- 29 January. 10 stars/points from the following problems:
  - at least 2 stars from 2.1-2.60 (different problems than you solved in the previous week)
  - at least 2 stars from 3.1-3.12
  - at least 2 stars from 3.13-3.24
- The problems will be given at the course web site

Presentations

- Timetable will be made today.
- Presentation should be on a given topic. However, please select what you are going to present appropriately (you can't go through all details in a given time). You can assume that people have (at least cursorily) read through the chapter, and that they know the basic math.
- Use of (also) whiteboard is encouraged.
- Duration of a presentation should be on average 45 minutes plus discussion.
- Hopefully there is discussion both during and after the talk.
- PDF's of the slides should be emailed to the assistant on previous Friday, at latest. They will be put to the course web site.

Chapter 1: Introduction

- Types of learning problems
- Probabilities (basic probabilities, expectations)
- Bayesian reasoning
- Regression (overfitting, regularization)
- Model selection (cross validation)
- (Curse of dimensionality)
- Information theory (what is information, KL divergence, mutual information)