Short course

A vademecum of statistical pattern recognition and machine learning

Introductory notes

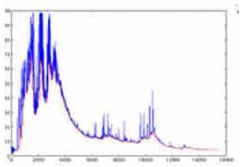
Massimo Piccardi
University of Technology, Sydney, Australia

What is statistical pattern recognition and why you should study it

- Recognising “patterns” of interest in data with the use of statistical techniques
- Unnumbered actual and potential applications:
  - image and video analysis
  - speech analysis
  - genomic research
  - biological data analysis
  - data mining and analytics for business, finance, marketing, trade
  - network traffic analysis
  - analysis of Web data
  - …

Example data

Australian Sign Language signs



Cancer detection from mass-spectrometric data

Chemical data to assess wine origin

O-ring data to predict thermal distress

examples courtesy of UCI Machine Learning Repository

Aims of this short course

- Offer a small “vademecum” (“carry with me”) of useful statistical pattern recognition techniques that you can later expand based on the need of your studies
- Minimal amount of proofs, informal presentation
- The selection of topics is oriented to ***learning and inference over sequential data***. It can also be a useful first step towards *structural data*
- Statistical pattern recognition is part of *machine learning* and is often assimilated to it

Topics at a glance

- Review of probability and statistics
- Density estimation
- Dimensionality reduction (with a preamble on linear regression)
- Bayesian classification
- The hidden Markov model
- Conditional random fields
- The support vector machine (SVM) and structural SVM
- Kalman and particle filters

Optimization

- Although during this short course we address maximization/minimization (i.e. optimization) of certain functions, we do not formally introduce optimization. You may like to refer to the videolectures and textbook of Prof. Stephen Boyd at Stanford:
  - http://www.stanford.edu/class/ee364a/videos.html
  - Stephen Boyd and Lieven Vandenberghe, Convex Optimization, Cambridge University Press, 2004, http://www.stanford.edu/~boyd/cvxbook/

Main reference textbooks

- Christopher M. Bishop, Pattern Recognition and Machine Learning, 2nd printing edition, Springer, 2007

  An excellent textbook, it has rapidly become the book of reference for the field

- David Barber, Bayesian Reasoning and Machine Learning, Cambridge University Press, 2012

  A very interesting, recent book which is freely available in its online version

- Ricardo Gutierrez-Osuna, Lecture Notes, http://research.cs.tamu.edu/prism/lectures.htm

  Excellent online materials; sets of widely commented, self-explanatory slides

- On Kalman and related topics, Yaakov Bar-Shalom's books; on particle filters, Arnaud Doucet *et al*.' book

