Machine Learning 2

0. Overview

Lars Schmidt-Thieme

Information Systems and Machine Learning Lab (ISMLL)
Institute for Computer Science
University of Hildesheim, Germany

Outline

1. Lecture Overview

2. Organizational Stuff

Outline

Syllabus

		A. Advanced Supervised Learning
Fri. 7.4.	(1)	A.1 Generalized Linear Models
Fri. 14.4.	—	— Good Friday —
Fri. 21.4.	(2)	A.2 Gaussian Processes
Fri. 28.4.	(3)	A.2b Gaussian Processes (ctd.)
Fri. 5.5.	(4)	A.3 Advanced Support Vector Machines
Fri. 12.5.	(5)	A.4 Neural Networks
Fri. 19.5.	(6)	A.5 Ensembles (Stacking)
Fri. 26.5.	(7)	A.5b Ensembles (Boosting, ctd.)
Fri. 2.6.	(8)	A.5c Ensembles (Mixtures of Experts, ctd.)
Fri. 9.6.	—	— Pentecoste Break —
Fri. 16.6.	(9)	A.6 Sparse Linear Models — L1 regularization
Fri. 23.6.	(10)	A.6b Sparse Linear Models — L1 regularization (ctd.)
Fri. 30.6.	(11)	A.7. Sparse Linear Models — Further Methods
		B. Complex Predictors
Fri. 7.7.	(12)	B.1 Latent Dirichlet Allocation (LDA)

Possible Further Topics

A. Advanced Supervised Learning
A.x Generalized Additive Models (Mur 16.3)

B. Complex Data (Relations, Images, Text, ...)
B.1 Statistical relational learning / Factorization models
B.2 Deep Learning / Representation Learning (Convolutional Neural Networks) (Mur. 28)

C. Complex Decisions
C.1 Ranking (Learning to rank) (Mur. 9.7)
C.2 Bayesian Regression & Classification (i.e., with uncertainties; variational methods; Gibbs sampling, MCMC) (Bishop 3.3, 4.5)
C.3 Sequential Classification (Conditional Random Fields/CRFs) (Mur. 19.6)
C.4 Structured Prediction

D. Problem Characteristics
D.1 Learning with additional unlabeled data (Semi-supervised Learning)
D.2 Controlling data acquisition (Active Learning)
D.3? Learning with missing values (imputation, EM)
D.4? Learning with imbalanced class distributions

E. Metalearning
E.1 Hyperparameter Learning

F. Learning theory
F.1 Bias/variance tradeoff; Union and Chernoff/Hoeffding bounds.
F.2 VC dimension
F.3 Problem Reductions

Outline

Character of the Lecture

This is an advanced lecture:

- I will assume good knowledge of Machine Learning I.
- Slides will contain major keywords, not the full story.
- For the full story, you need to read the referenced chapters in one of the books.

Exercises and Tutorials

- There will be a weekly sheet with 2 exercises handed out **each Wednesday** in the lecture.
  1st sheet will be handed out next week, Wed. 12.4.
- Solutions to the exercises can be submitted until **next Sunday 23:59 pm**
  1st sheet is due Sun. 16.4., 23:59 pm.
- Exercises will be corrected.
- Tutorials **each Wednesday 2pm–4pm**,
  1st tutorial next week, Wed. 12.4.
- Successful participation in the tutorial gives up to 10% bonus points for the exam.

Exam and Credit Points

- There will be a written exam at end of term (2h, 4 problems).
- The course gives 6 ECTS (2+2 SWS).
- The course can be used in
  - IMIT MSc. / Informatik / Gebiet KI & ML
  - Wirtschaftsinformatik MSc / Informatik / Gebiet KI & ML & Wirtschaftsinformatik MSc / Wirtschaftsinformatik / Gebiet BI
  - as well as in both BSc programs.

Some Books

- Kevin P. Murphy (2012):
  Machine Learning, A Probabilistic Approach, MIT Press.
- Trevor Hastie, Robert Tibshirani, Jerome Friedman (^{2}2009):
  The Elements of Statistical Learning, Springer.

  Also available online as PDF at `http://www-stat.stanford.edu/~tibs/ElemStatLearn/`
- Christopher M. Bishop (2007):
  Pattern Recognition and Machine Learning, Springer.
- Richard O. Duda, Peter E. Hart, David G. Stork (^{2}2001):
  Pattern Classification, Springer.

Further Readings

- For a general introduction: [JWHT13, chapter 1&2], [Mur12, chapter 1], [HTFF05, chapter 1&2].
- For linear regression: [JWHT13, chapter 3], [Mur12, chapter 7], [HTFF05, chapter 3].

References

Trevor Hastie, Robert Tibshirani, Jerome Friedman, and James Franklin.
The elements of statistical learning: data mining, inference and prediction, volume 27.
Springer, 2005.

Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani.
An introduction to statistical learning.
Springer, 2013.

Kevin P. Murphy.
Machine learning: a probabilistic perspective.
The MIT Press, 2012.