# 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. | 24.4. | (1) | A.1 Generalized Linear Models |
| Fri. | 1.5. | — | — *Labour Day* — |
| Fri. | 8.5. | (2) | A.2 Gaussian Processes |
| Fri. | 15.5. | (3) | A.3 Advanced Support Vector Machines |
| | | | **B. Ensembles** |
| Fri. | 22.5. | (4) | B.1 Stacking |
| | | | & B.2 Boosting |
| Fri. | 29.5. | (5) | B.3 Mixtures of Experts |
| Fri. | 5.6. | — | — *Pentecoste Break* — |
| | | | **C. Sparse Models** |
| Fri. | 12.6. | (6) | C.1 Homotopy and Least Angle Regression |
| Fri. | 19.6. | (7) | C.2 Proximal Gradients |
| Fri. | 26.6. | (8) | C.3 Laplace Priors |
| Fri. | 3.7. | (9) | C.4 Automatic Relevance Determination |
| | | | **D. Complex Predictors** |
| Fri. | 10.7. | (10) | D.1 Latent Dirichlet Allocation (LDA) |
| Fri. | 17.7. | (11) | Q & A |

# 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

- Two tutorials (course 3102):
  - Tuesdays, 8:00-10:00 in A009
  - Wednesdays, 14:00-16:00 in C213
- Submit your preferences until next week in the LearnWeb Poll
- Tutorial sheet upload: each **Thursday, 10:00 AM** (LearnWeb)
- Tutorial sheet deadline: **Thursday, 10:00 AM**
  (upload or boxes at Samelsonplatz)
- Earn up to 10% bonus points for the exam.
- Next week: in class tutorial.

# 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: [**?**, chapter 1&2], [**?**, chapter 1], [**?**, chapter 1&2].
- For linear regression: [**?**, chapter 3], [**?**, chapter 7], [**?**, 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.