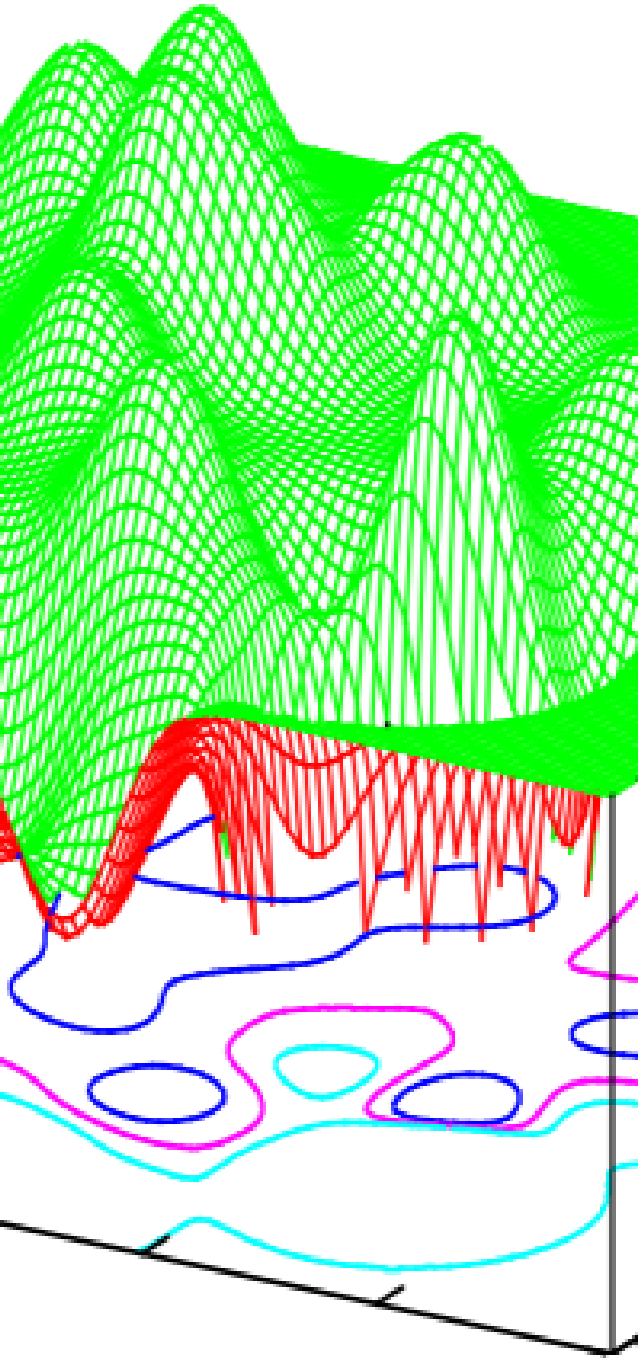


Machine Learning

Introduction

Marc Toussaint
University of Stuttgart
Summer 2016

What is Machine Learning?

1) A long list of methods/algorithms for different data anlysis problems
   - in sciences
   - in commerce

2) Frameworks to develop your own learning algorithm/method

3) Machine Learning = information theory/statistics + computer science

What is Machine Learning?

- Pedro Domingos: *A Few Useful Things to Know about Machine Learning*

LEARNING = REPRESENTATION + EVALUATION + OPTIMIZATION

- "Representation": Choice of model, choice of hypothesis space
- "Evaluation": Choice of objective function, optimality principle

  Notes: The *prior* is both, a choice of representation and, usually, a part of the objective function.
  In Bayesian settings, the choice of model often directly implies also the "objective function"
- "Optimization": The algorithm to compute/approximate the best model

Pedro Domingos: *A Few Useful Things to Know about Machine Learning*

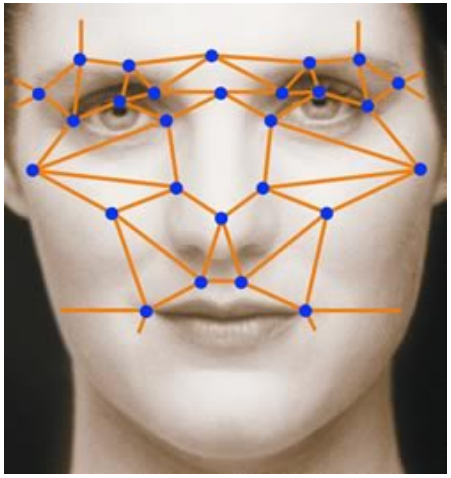
- It's generalization that counts
  - Data alone is not enough
  - Overfitting has many faces
  - Intuition fails in high dimensions
  - Theoretical guarantees are not what they seem
- Feature engineering is the key
- More data beats a cleverer algorithm
- Learn many models, not just one
- Simplicity does not imply accuracy
- Representable does not imply learnable
- Correlation does not imply causation

Machine Learning is everywhere

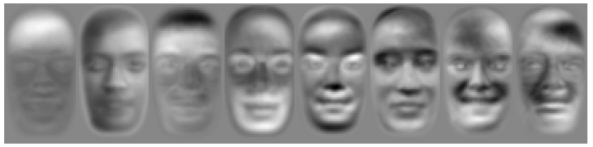
NSA, Amazon, Google, Zalando, Trading, ...
Chemistry, Biology, Physics, ...
Control, Operations Reserach, Scheduling, ...

- Machine Learning $\sim$ Information Processing (e.g. Bayesian ML)

Face recognition



keypoints



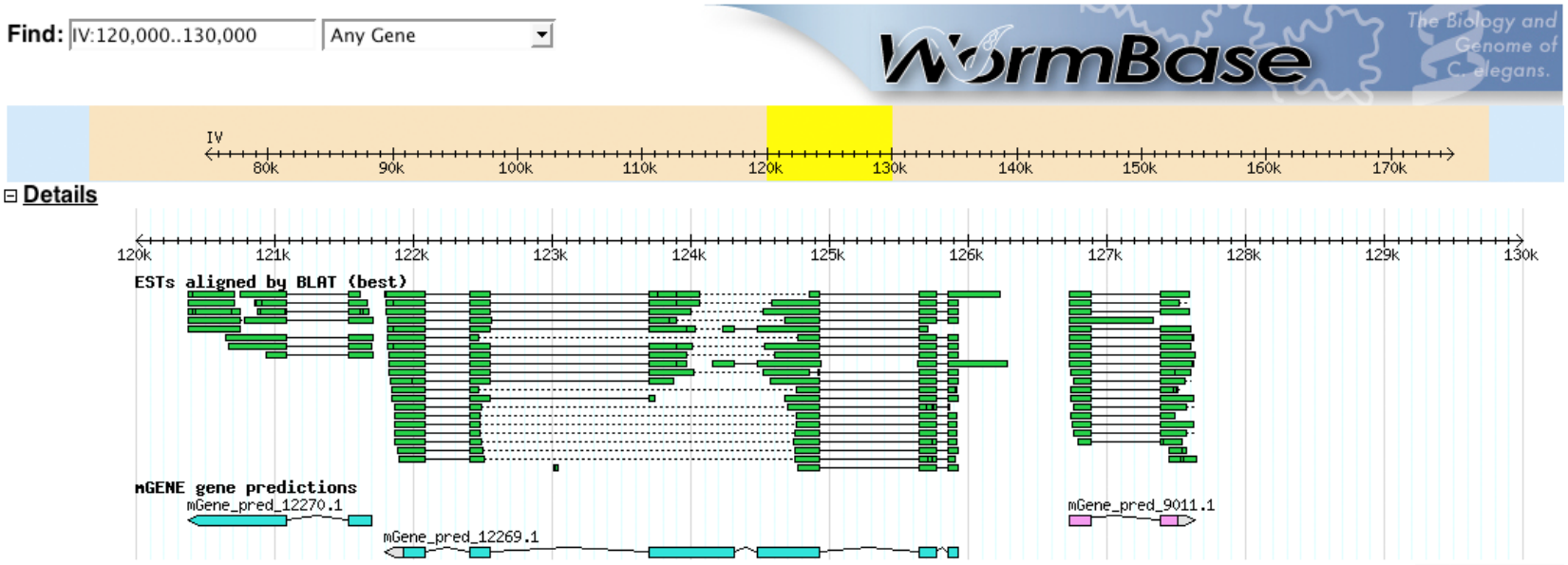
eigenfaces

(e.g., Viola & Jones)

Hand-written digit recognition (US postal data)

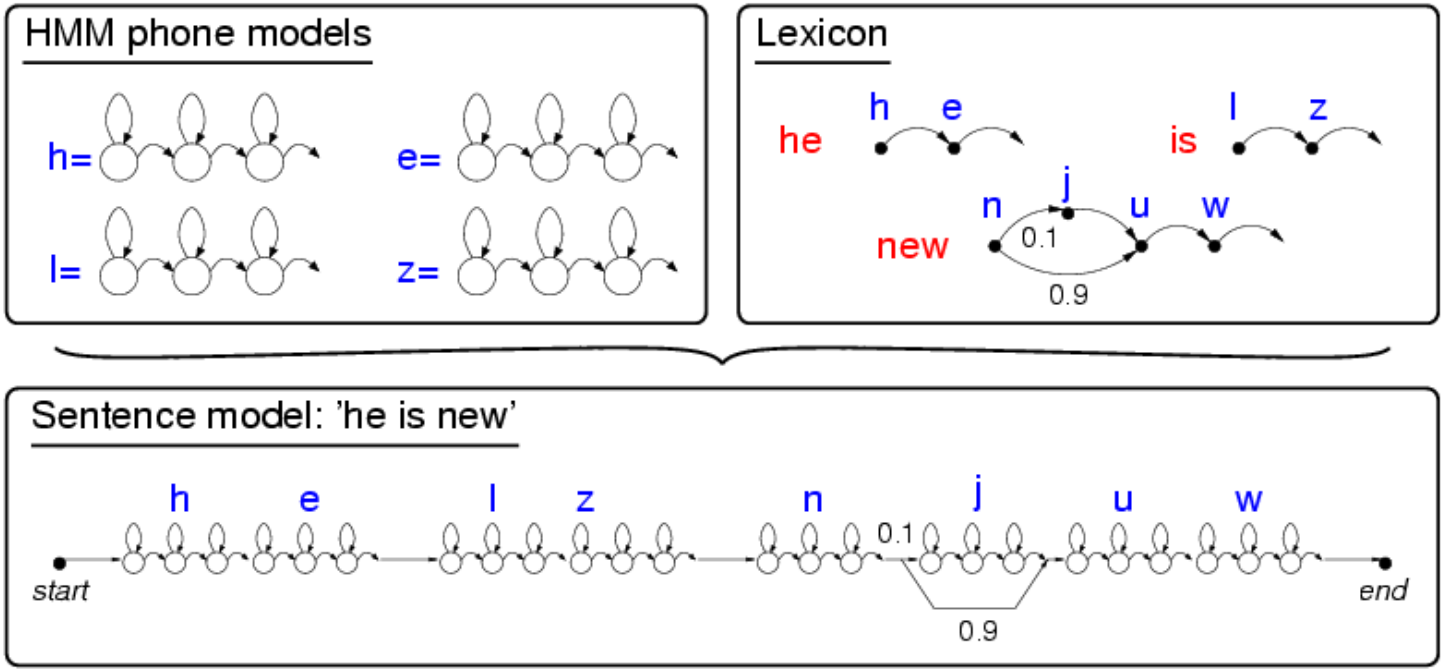
(e.g., Yann LeCun)

Gene annotation



(Gunnar Rätsch, Tübingen, mGene Project)

Speech recognition



(This is the basis of all commercial products)

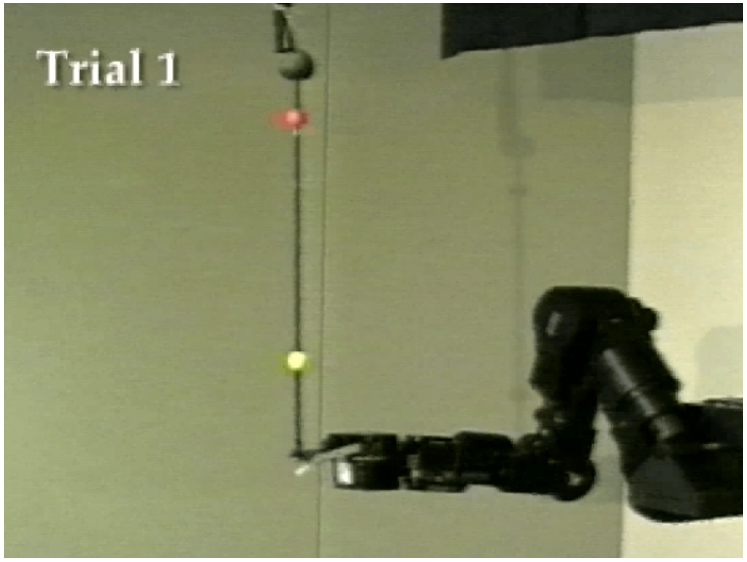
Spam filters

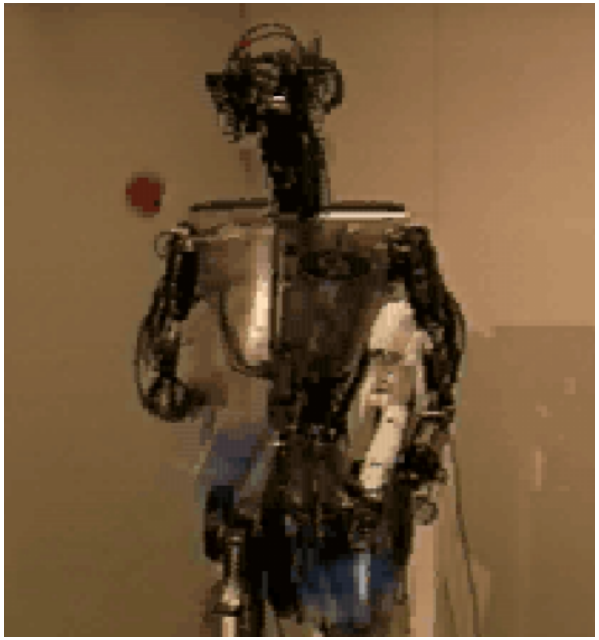
	george	you	your	hp	free	hpl	!	our	re	edu	remove
spam	0.00	2.26	1.38	0.02	0.52	0.01	0.51	0.51	0.13	0.01	0.28
email	1.27	1.27	0.44	0.90	0.07	0.43	0.11	0.18	0.42	0.29	0.01

Machine Learning became an important technology
in science as well as commerce

Examples of ML *for behavior*...

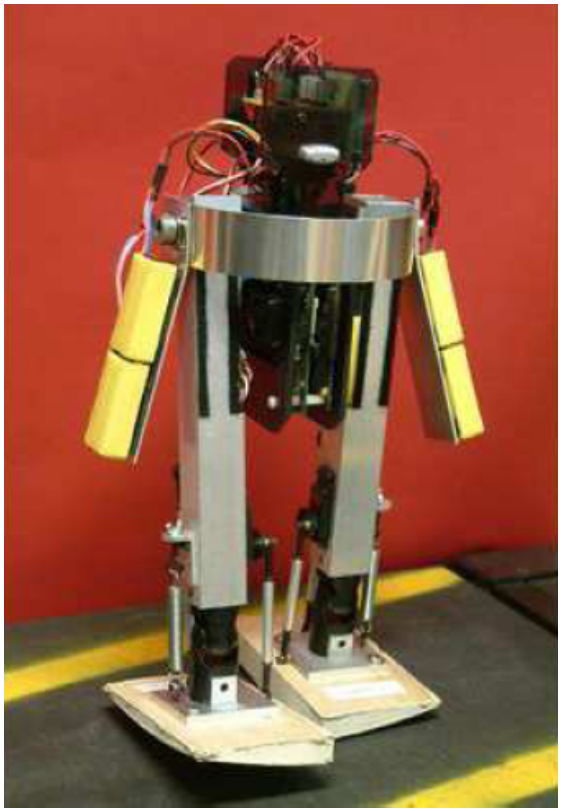
Learing motor skills





(around 2000, by Schaal, Atkeson, Vijayakumar)

Learning to walk



(Rus Tedrake et al.)

Organization of this lecture

See TOC of last year's slide collection

- Part 1: The Core: Regression & Classification
- Part 2: The Breadth of ML methods
- Part 3: Bayesian Methods

Is this a theoretical or practical course?

Neither alone.

Is this a theoretical or practical course?

Neither alone.

- The goal is to teach how to design good learning algorithms

data

$\downarrow$

modelling [requires theory & practise]

$\downarrow$

algorithms [requires practise & theory]

$\downarrow$

testing, problem identification, restart

How much math do you need?

- Let $L(x) = \|y - Ax\|^2$. What is

$$\operatorname*{argmin}_x L(x)$$

- Find

$$\min_x \|y - Ax\|^2 \quad \text{s.t.} \quad x_i \leq 1$$

- Given a discriminative function $f(x, y)$ we define

$$p(y \mid x) = \frac{e^{f(y,x)}}{\sum_{y'} e^{f(y',x)}}$$

- Let A be the covariance matrix of a Gaussian. What does the Singular Value Decomposition $A = VDV^\top$ tell us?

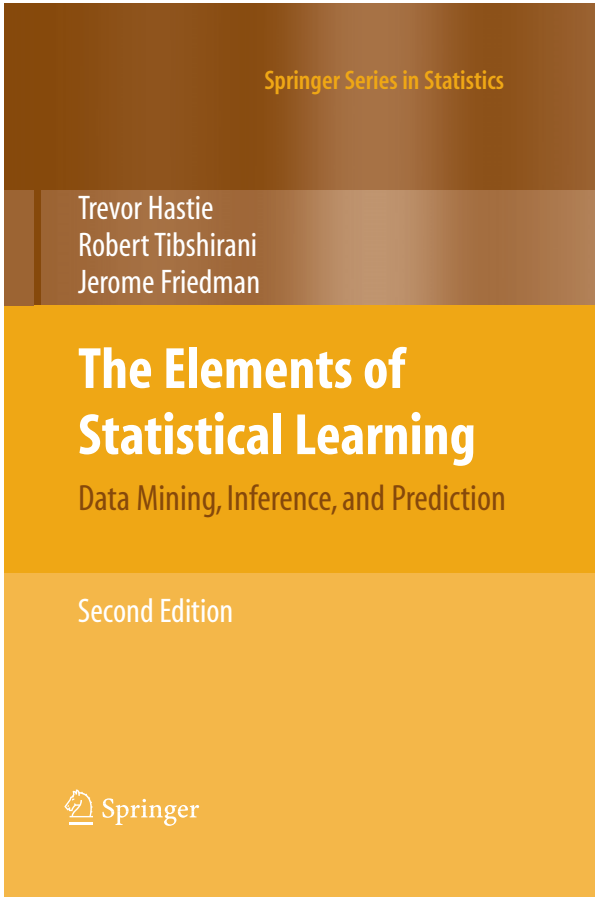
How much coding do you need?

- A core subject of this lecture: learning to go from principles (math) to code

- Many exercises will implement algorithms we derived in the lecture and collect experience on small data sets

- Choice of language is fully free. I support C++; tutors might prefer Python; Octave/Matlab or R is also good choice.

Books



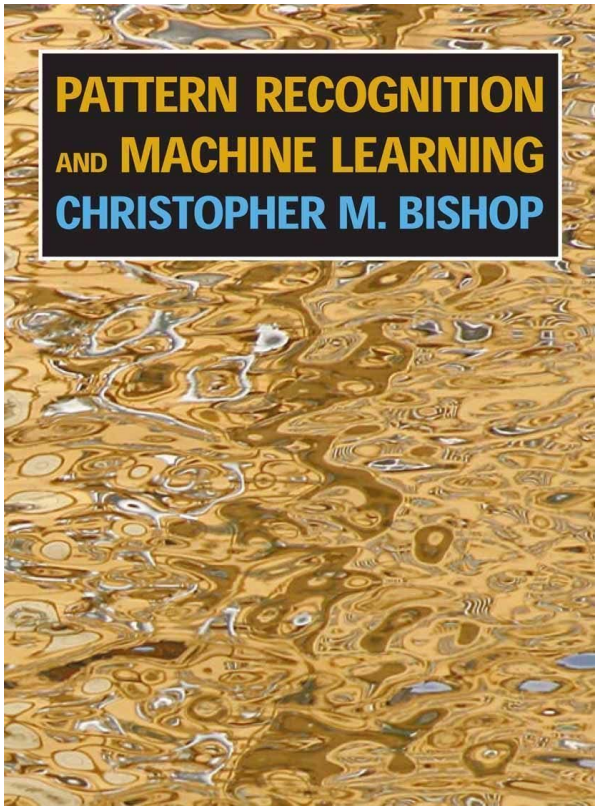
Trevor Hastie, Robert Tibshirani and Jerome Friedman: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* Springer, Second Edition, 2009.

`http://www-stat.stanford.edu/~tibs/ElemStatLearn/`

(recommended: read introductory chapter)

(this course will not go to the full depth in math of Hastie et al.)

Books



Bishop, C. M.: *Pattern Recognition and Machine Learning*.
Springer, 2006
`http://research.microsoft.com/en-us/um/people/cmbishop/prml/`
(some chapters are fully online)

Books & Readings

- more recently:
  - David Barber: Bayesian Reasoning and Machine Learning
  - Kevin Murphy: Machine learning: a Probabilistic Perspective

- See the readings at the bottom of:

`http://ipvs.informatik.uni-stuttgart.de/mlr/marc/teaching/index.html#readings`

Organization

- Course Webpage:

`http://ipvs.informatik.uni-stuttgart.de/mlr/marc/teaching/16-MachineLearning/`

  - Slides, Exercises & Software (C++)
  - Links to books and other resources

- Admin things, please first ask:
  Carola Stahl, `Carola.Stahl@ipvs.uni-stuttgart.de`, Raum 2.217

- Rules for the tutorials:
  - Doing the exercises is crucial!
  - **Nur Votieraufgaben.** At the beginning of each tutorial:
    – sign into a list
    – mark which exercises you have (successfully) worked on
  - Students are randomly selected to present their solutions
  - You need 50% of completed exercises to be allowed to the exam
  - Please check 2 weeks before the end of the term, if you can take the exam