DM825 (5 ECTS - 3rd Quarter)
Introduction to Machine Learning
[Introduktion til maskinlæring]

Marco Chiarandini
lektor, IMADA
www.imada.sdu.dk/~marco

Machine Learning

Machine learning is the science of getting computers to act without being explicitly programmed.

Application examples:

- practical speech recognition and automatic machine translation
- effective web search
- email spam detection
- social networks recognize friends from photos or suggest friends
- understanding of the human genome
- ...

The course focuses on the theoretical background and the practical application of a varied group of techniques.

Contents

- Supervised Learning
  - Classification and Regression via Linear Models
  - Neural Networks
  - Support Vector Machines
  - Probabilistic Graphical Models
    Bayesian Networks, Hidden Markov Models
  - Assessment and Selection
- Unsupervised Learning
  - Mixture Models and Expectation Maximization
  - Association rules, cluster analysis, principal components, frequent pattern mining

Example: Medical diagnosis

- two diseases: Flu and Hayfever
- they are not mutually exclusive
- season might be correlated with them
- symptoms such as Congestion and Muscle Pain

Interrelated aspects of the system are modeled as random variables:

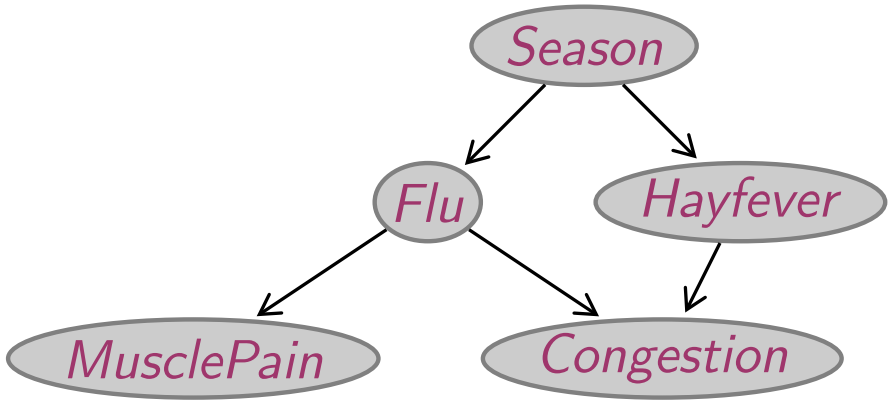
$Flu = \{\texttt{true}, \texttt{false}\}$
$Hayfever = \{\texttt{true}, \texttt{false}\}$
$Season = \{\texttt{fall}, \texttt{winter}, \texttt{spring}, \texttt{summer}\}$
$Congestion = \{\texttt{true}, \texttt{false}\}$
$MusclePain = \{\texttt{true}, \texttt{false}\}$

$2 \times 2 \times 4 \times 2 \times 2 = 64$
possible prob. values
for joint distribution

$$P(Flu = \texttt{true} \mid Seas. = \texttt{fall}, Cong. = \texttt{true}, MusclePain = \texttt{false}) = ?$$

Graphical Models

Encode uncertainty and our prior knowledge in a graphical model



F and H independent given Season
C and S independent given F and H
M and H,C independent given F
M and C independent given F

We thus only need to define
$3 + 4 + 4 + 4 + 2 = 17$ parameters

$$P(S, F, H, C, M) = P(S)P(F \mid S)P(H \mid S)P(C \mid F, H)P(M \mid F)$$

Learning

What can we do from here?

- Inference: Complexity issues $O(2^n)$
- Learning (parameters and structure)

Learning Example: Coin Experiment
Flip coin and observe the number of times it lands with head and tail.
We wish to learn how much the probability deviates from 0.5.

Suppose we observe 3 heads in 10 tosses.

- With no prior knowledge we would set $p = 3/10 = 0.33$
- With a prior of 10 heads over 20 tosses we would set $p = (3 + 10)/(10 + 20) = 13/30 = 0.43$
- However if we obtain more data the effect diminishes:
  $(300 + 10)/(1000 + 20) = 0.3$
  $(300 + 1)/(1000 + 2) = 0.3$.

Aims of the Course

After the course you should be able to:

- recognize which learning method is suitable for a given task
- describe the theory behind the methods
- apply the method to example problems with few data
- undertake an experimental assessment of learning methods and report the results

Course Information (1/2)

Prerequisites:

- ✓ DM527 Mathematical Tools,
- ✓ MM501/MM502 Calculus I and II,
- ✓ DM502/DM503 Programming A and B,
- ✓ MM505 Linear algebra
- – ST501/ST502 Science Statistics/Statistical Modeling are an asset

Evaluation:

(a) Two mandatory assignments, pass/fail, internal evaluation by the teacher. Applied character, include programming work (in R)

(b) 3 hours written exam, Danish 7-point grading scale, External censorship.

Course Information (2/2)

Format:	2×2h per week=28h lectures + 8/10h exercises
Language:	English
Material:	▶ No text book ▶ M.C. Bishop, *Pattern Recognition and Machine Learning* 1st ed. 2006. Springer. 305 pp. ▶ Andrew Ng. Lecture Notes. `cs229.standford.edu/materials.html` ▶ Lecture notes in form of slides ▶ Articles
Credits & Term:	5 ECTS, 3rd Quarter

DM825 (5 ECTS - 3rd Quarter)
Introduction to Machine Learning
[Introduktion til maskinlæring]

Marco Chiarandini
lektor, IMADA
www.imada.sdu.dk/~marco