Learning Systems

Antanas Verikas
antanas.verikas@hh.se

IDE, Halmstad University

2013

Outline

Content

- Machine learning algorithms
- Problems attacked with learning systems
  - Classification
  - Regression
- Main issues in learning
  - Model bias and variance,
  - Over-fitting,
  - Generalization
- Practical use of machine learning

Main issues in learning

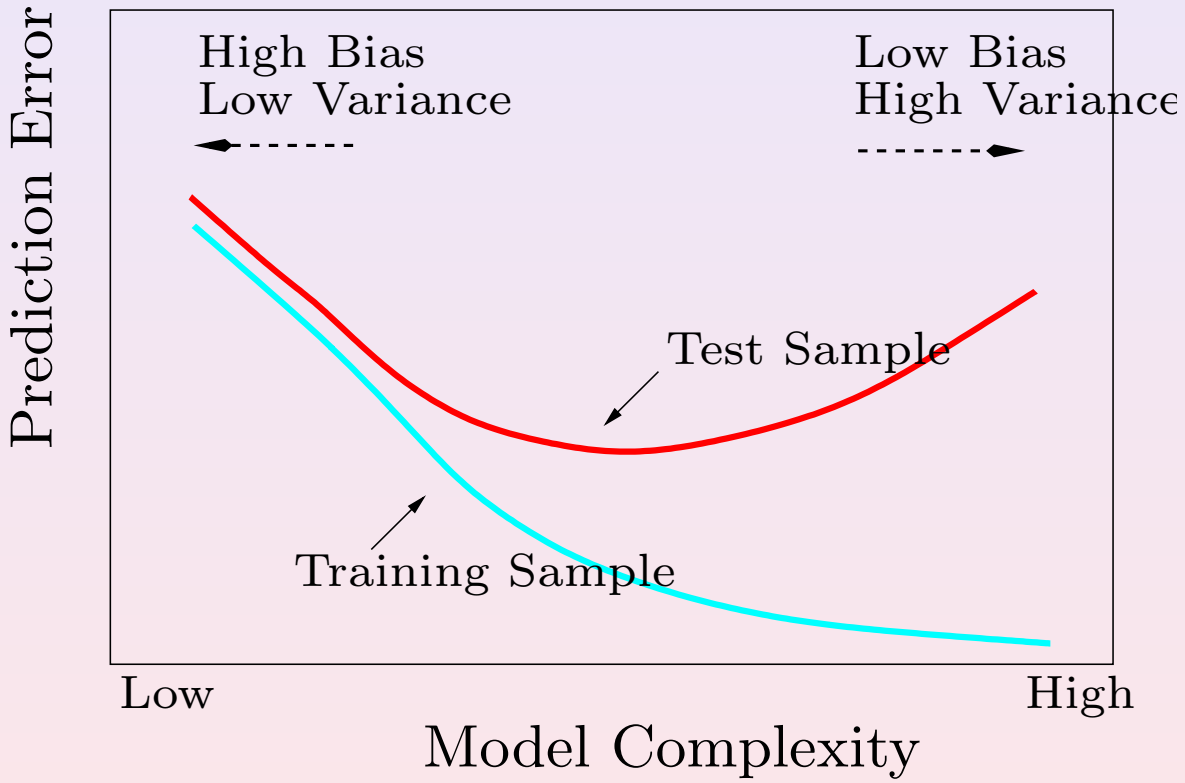


Figure: Relation between the prediction error and model complexity.

Goal

After taking the course you should get:

- Basic knowledge of the common linear and some nonlinear machine learning algorithms.
- A good understanding of neural network algorithms.

After taking the course you should be able to:

- judge when ML is applicable
- apply ML methods on real-world problems
- assimilate scientific papers in the ML area

Format

- Lectures (2 hrs/occasion)
  - Oral examination (40%)
- Two projects
  - Written report & oral presentation (40%)
- Seminars (1 each student)
  - Evaluation (20%)

Projects

You will do projects in groups of two students. Each group shall do two projects, one regression and one classification project. The projects will be performed in MATLAB. Results are to be presented in a report and orally.

Seminars

Each student has to give a seminar (15 mins). Seminars will be on applications, typically based on research papers. Material is provided by the course responsible.

Preliminary schedule

Preliminary schedule

Week 14				
T	4 Apr	08:15-10:00	Lecture	D415
F	5 Apr	10:15-12:00	Lecture	R3149
Week 15				
M	8 Apr	08:15-10:00	Lecture	D415
W	10 Apr	08:15-10:00	Lecture	D415
Week 16				
M	15 Apr	08:15-10:00	Lecture	D415
T	16 Apr	13:15-15:00	Lecture	D215
Week 17				
M	22 Apr	08:15-10:00	Lecture	D415
T	23 Apr	08:15-10:00	Lecture	D415

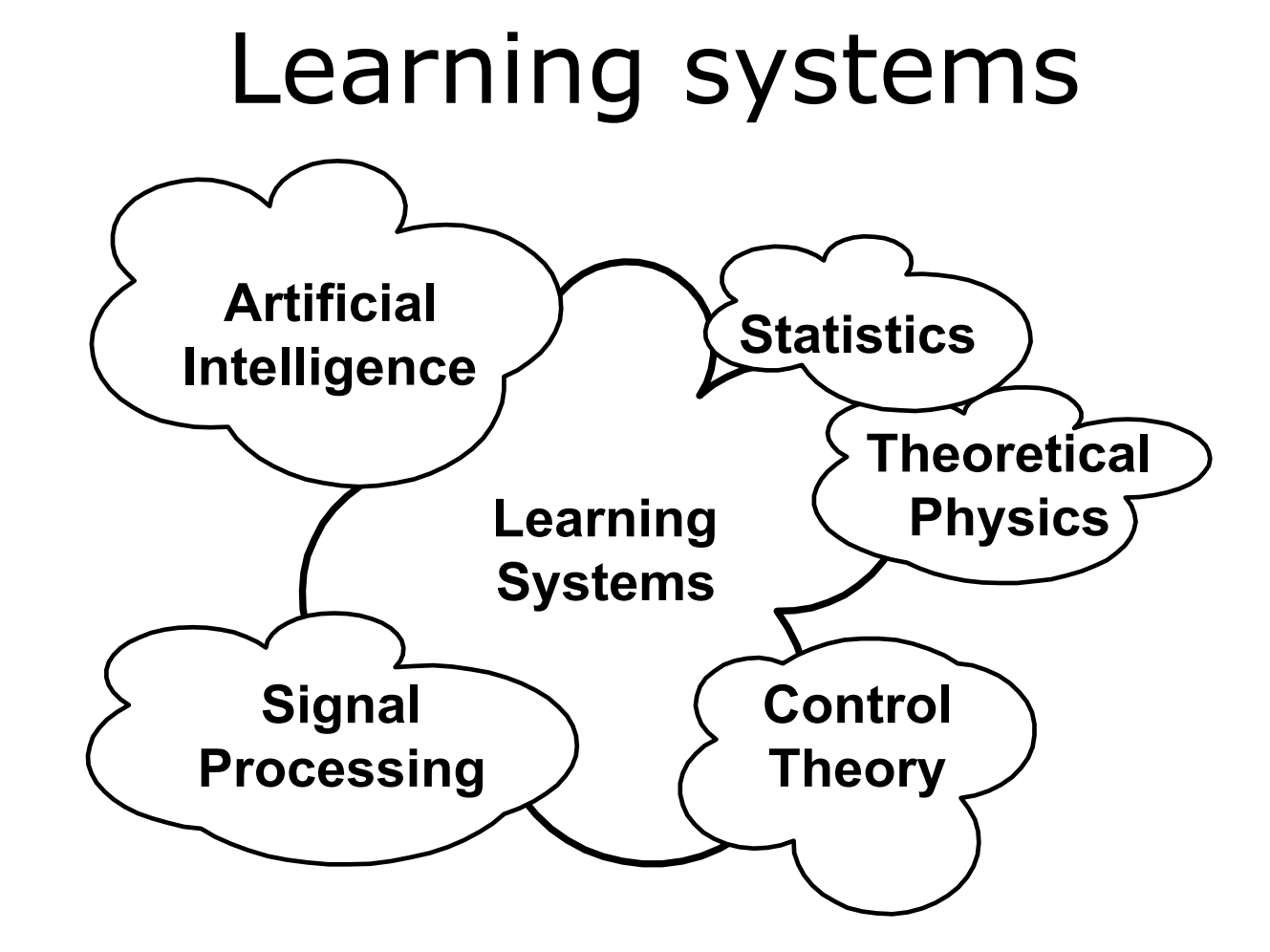
Preliminary schedule

Week 18				
M	29 Apr	08:15-10:00	Lecture	D415
T	30 Apr	08:15-10:00	Lecture	D415
Week 19				
M	06 May	08:15-10:00	Seminar	D415
T	07 May	08:15-10:00	Seminar	D415
Week 20				
M	13 May	08:15-10:00	Seminar	D415
Week 21				
M	20 May	08:15-10:00	Project	D415
T	21 May	08:15-10:00	Project	D415

Books

1. T. Hastie, R. Tibshirani, J. H. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction (Springer Series in Statistics), 2nd Edition, Springer-Verlag, New York, 2009. http://www.stanford.edu/ hastie /local.ftp/Springer/ESLII print5.pdf
2. R. O. Duda, P. E. Hart, D. G. Stork, Pattern Classification, 2nd Edition, John Wiley & Sons, New York, 2001.

Links with other areas



Intelligent Control

Intelligent Control

"An intelligent system must be highly adaptable to significant unanticipated changes, and so learning is essential. It must exhibit a high degree of autonomy in dealing with changes. It must be able to deal with significant complexity..."

Panos Antsaklis et al. (1994)

Artificial Intelligence

Artificial Intelligence (AI)

"...is concerned with intelligent behavior in artifacts. Intelligent behavior, in turn, involves perception, reasoning, learning, communicating, and acting in complex environments."

Nils J. Nilsson, 1998

Artificial Intelligence

Problem areas

- Game playing
- Automated reasoning
- Spoken language interpretation
- Autonomous robots
- Expert systems
- Machine learning

Expert Systems

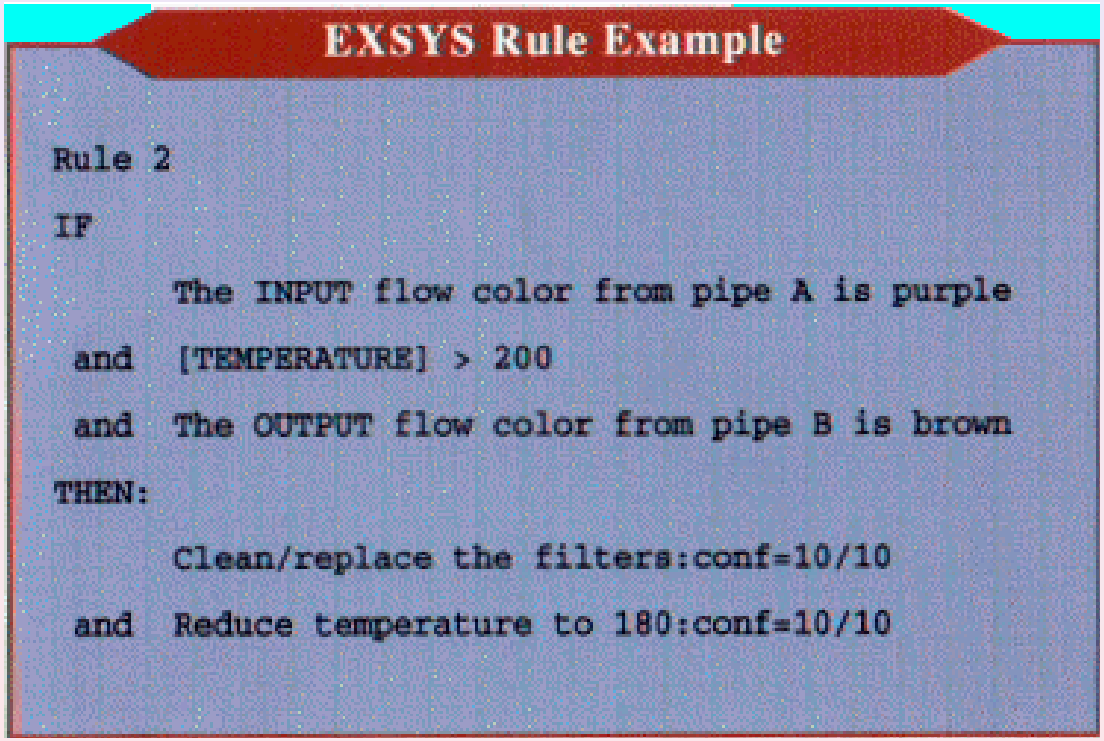


Figure: Image from Apexs, Inc

- Based on IF...THEN rules
- Usually requires human expert

Types of learning in AI

- *Deductive*: Deduce rules from already known rules (Example: Mathematics)

$$(A \Rightarrow B \Rightarrow C) \Rightarrow (A \Rightarrow C)$$

- *Inductive*: Learn new rules from data

$$\mathbf{Z} = \{\mathbf{x}(n), y(n)\}_{n=1,\ldots,N} \Rightarrow (A \Rightarrow C)$$

We will only be dealing with *inductive* learning.

Two types of inductive learning

- *Supervised*: The machine has access to a teacher who corrects it.
- *Unsupervised*: No access to a teacher. The machine must search for "order" in the environment.

We will study both, but mainly supervised.

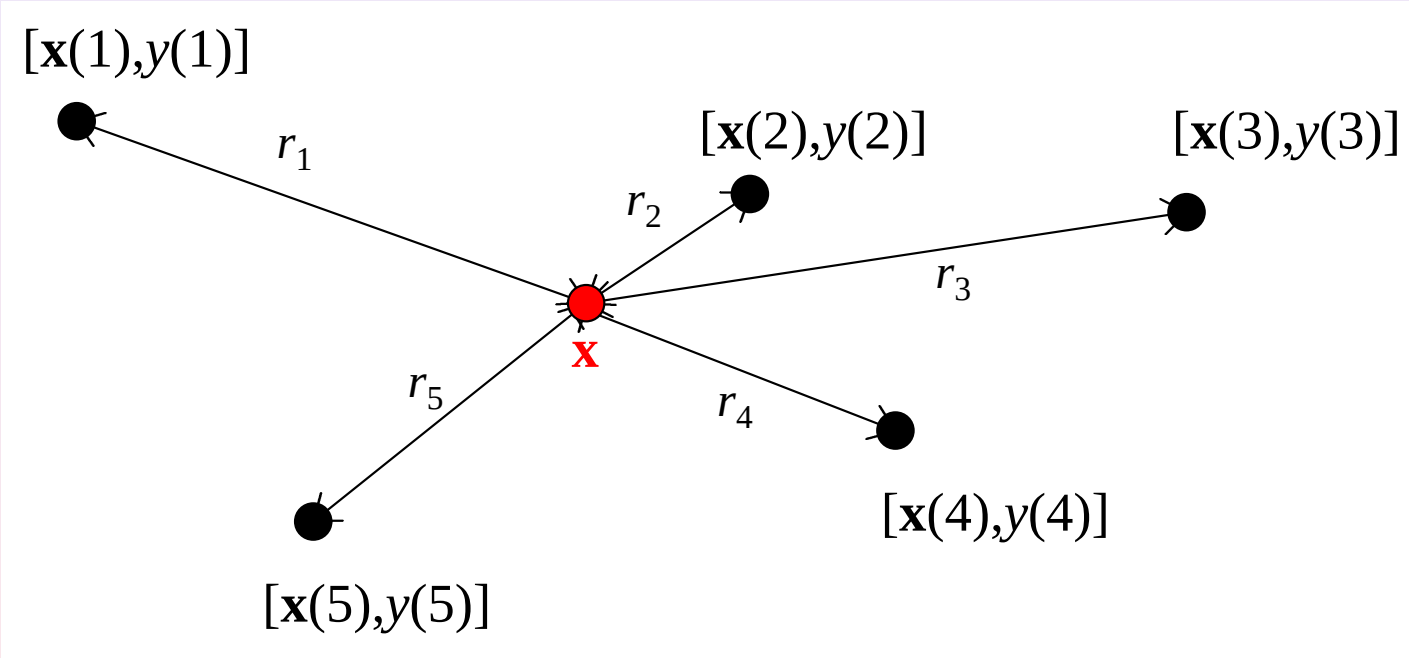
Supervised learning

- *Memory based*: There is no model, just a database is built. Example: k-Nearest Neighbours (k-NN).
- *A model is built based on error correction*: Minimize error E:

$$E(\mathbf{w}) = \sum_{n=1}^{N} \{f[\mathbf{w}, \mathbf{x}(n)] - y(n)\}^2$$

where $\mathbf{w}$ are the model's parameters.

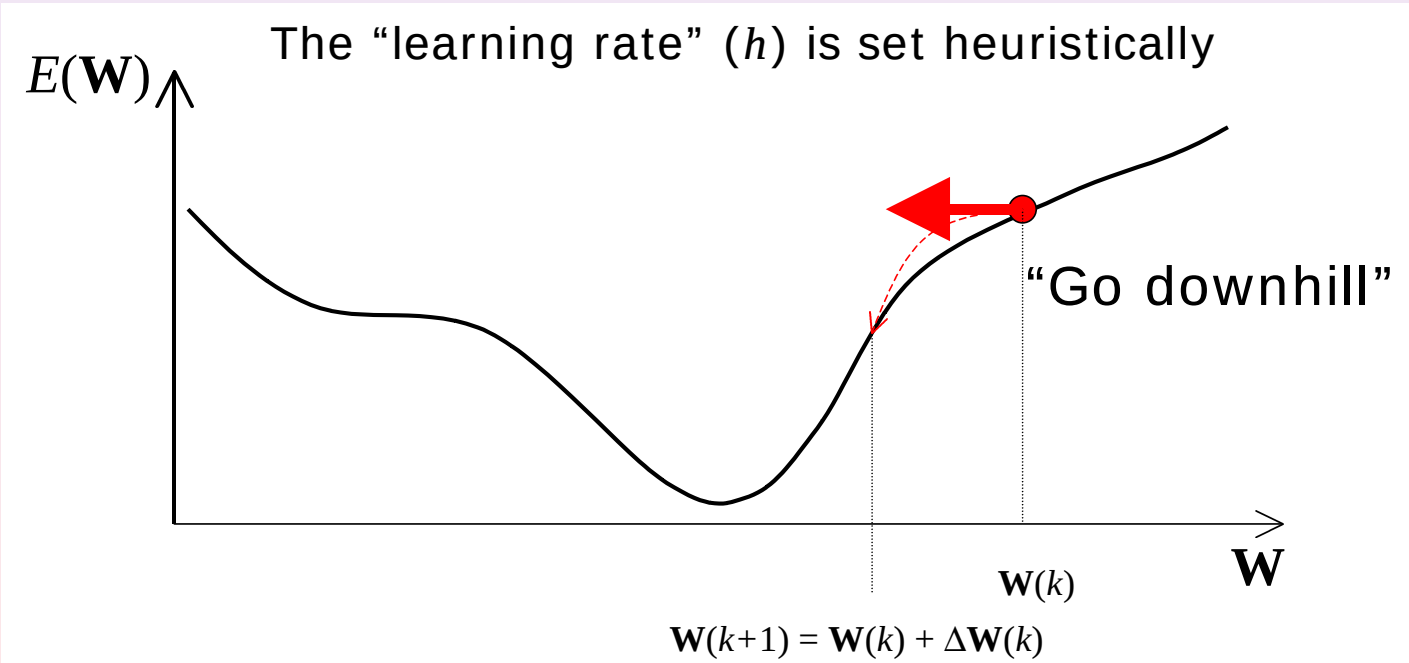
k-NN type algorithm



$$r_2 < r_j, \quad \forall j \neq 2 \Rightarrow f(k, \mathbf{x}) = y(2)$$

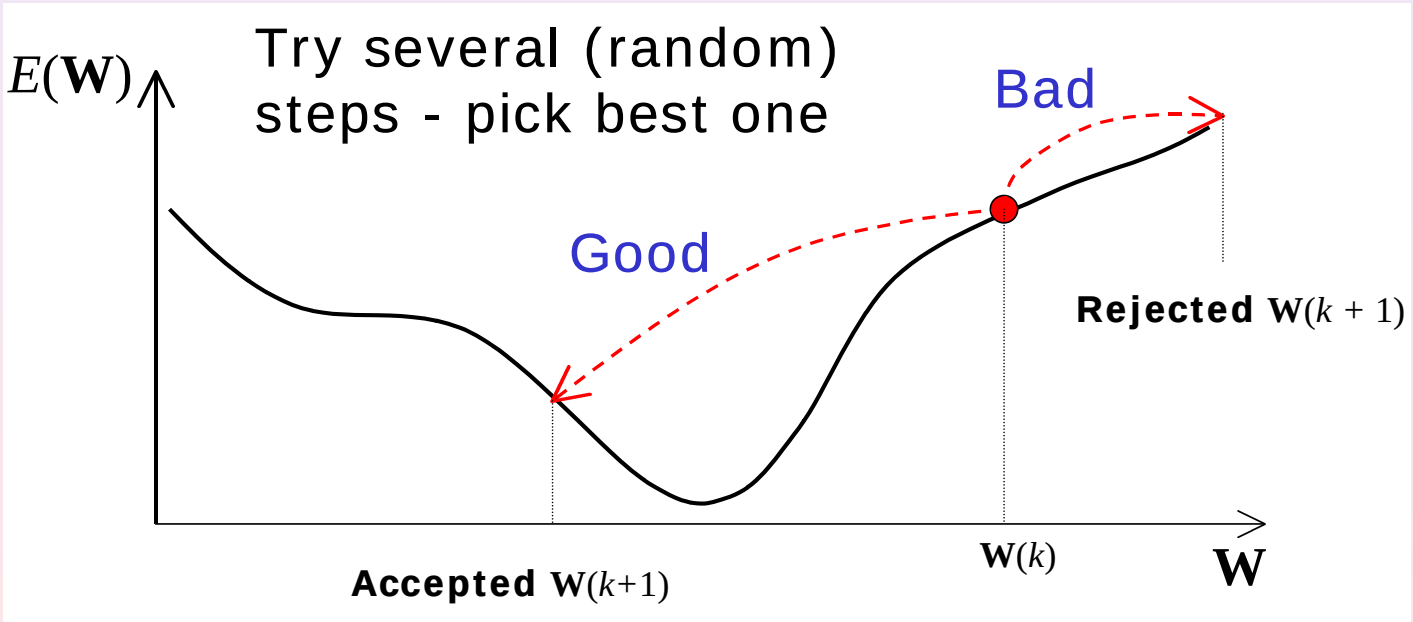
Error correction (gradient-based)

$$\Delta \mathbf{w} = -\eta \nabla_{\mathbf{w}} E(\mathbf{w})$$



Error correction ("random" search-based)

Is used if ∇E is impossible or difficult to compute, but $E(\mathbf{w})$ is easy to compute.



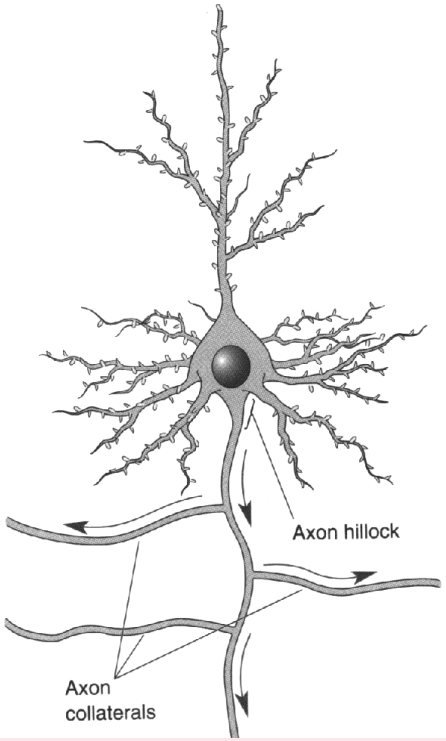
Why Machine Learning?

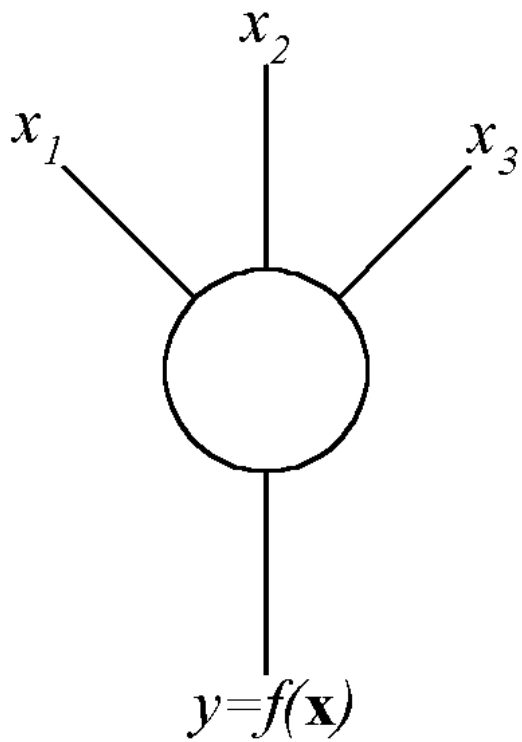
- Some tasks can only be described by examples.
- Need for on-line tuning.
- Too difficult to write down all details.
- Knowledge increases, too difficult to update by hand.

Some Machine Learning Methods

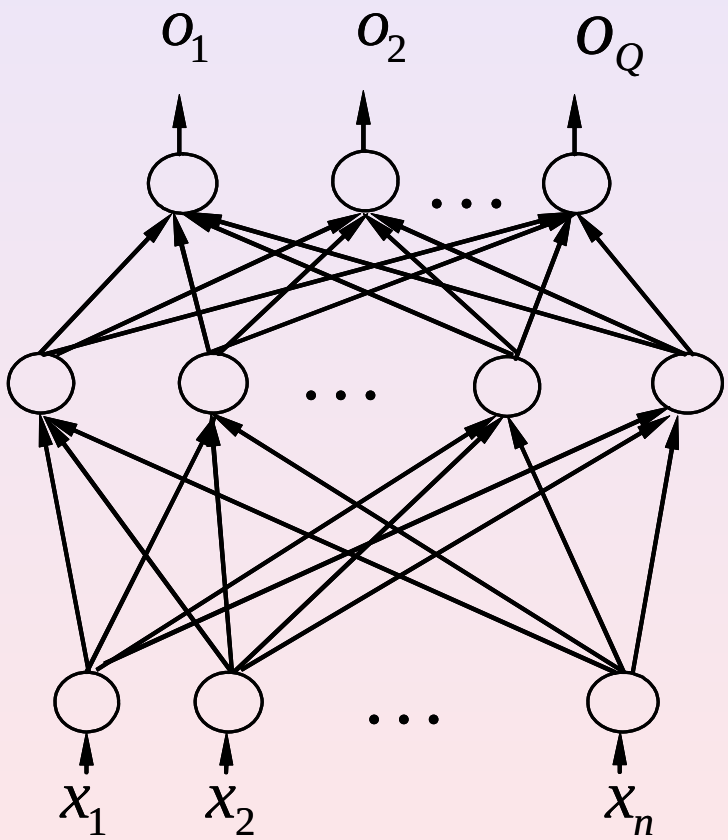
- Artificial neural networks (ANN)
  - Models inspired by the structure of the neural system.
- Decision trees
  - Similar to expert systems, produce rules.
- Genetic algorithms
  - Mimics how life evolves via mutations, mating, and "natural selection".
- Support vector machines (SVM)

Biological and Artificial Neuron





Artificial Neural Network



ANN in Apple Newton

Figure: Apple Newton.

- Relatively low accuracy of handwritten character recognition.
- ANN give much better performance.
- "...vastly improved handwriting recognition..." (BYTE May 1996)
- In production 1995.

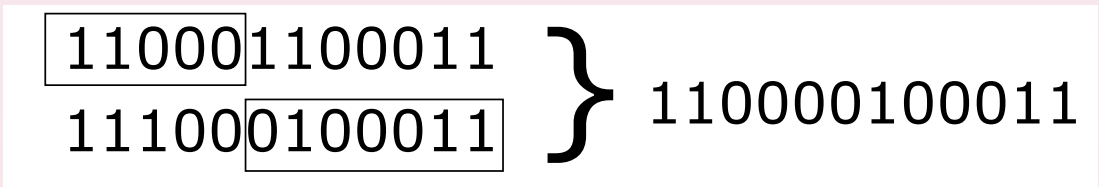
ANN in Foss Tecator Grain Checker

Figure: Foss Tecator Grain Checker.

- Analyzes quality of grain from an image.

Genetic algorithms

- Represent probable solutions as binary strings
  - solution A: 110001100011
  - solution B: 110011100101
  - solution C: 111000100011
  - solution ...
- Select "parent solutions" based on how "good" they are. Let them mate: (e.g. A+C)



Crossover and mutation

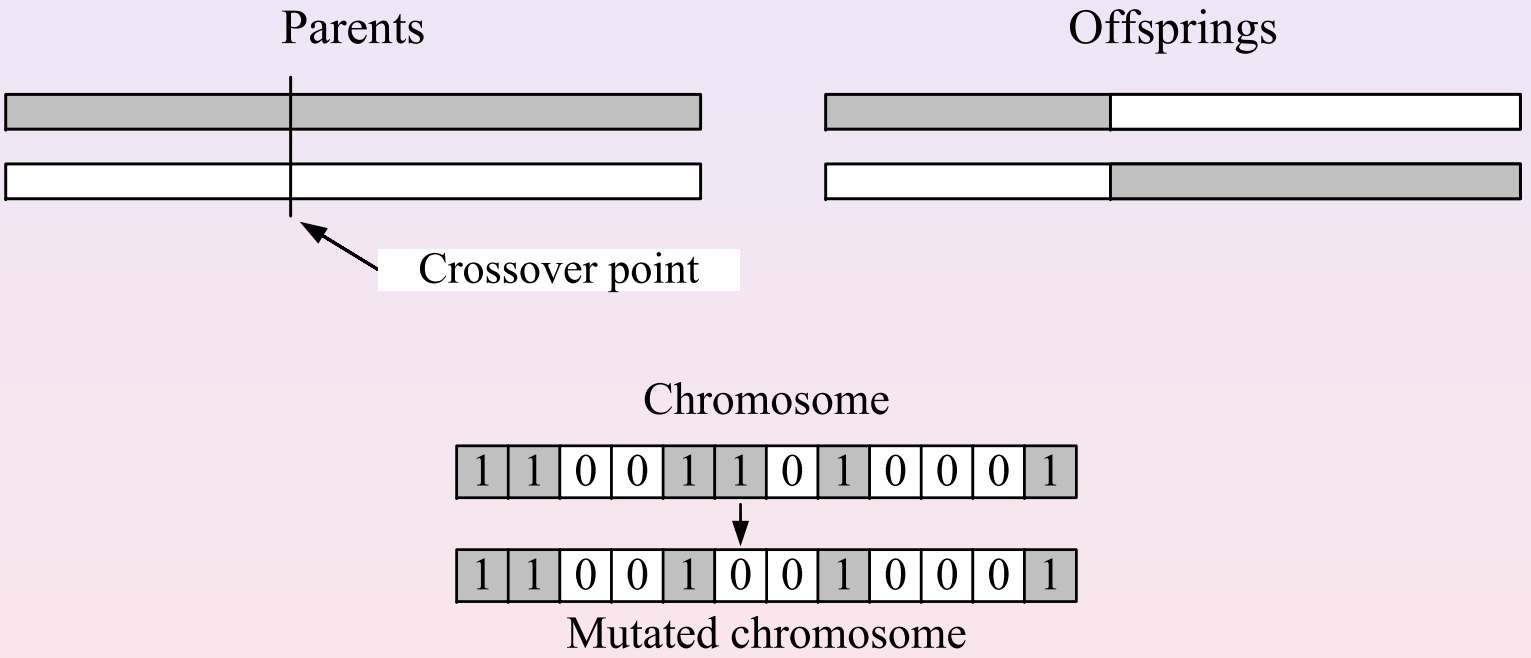


Figure: Crossover (*top*) and mutation (*bottom*).

Genetic search

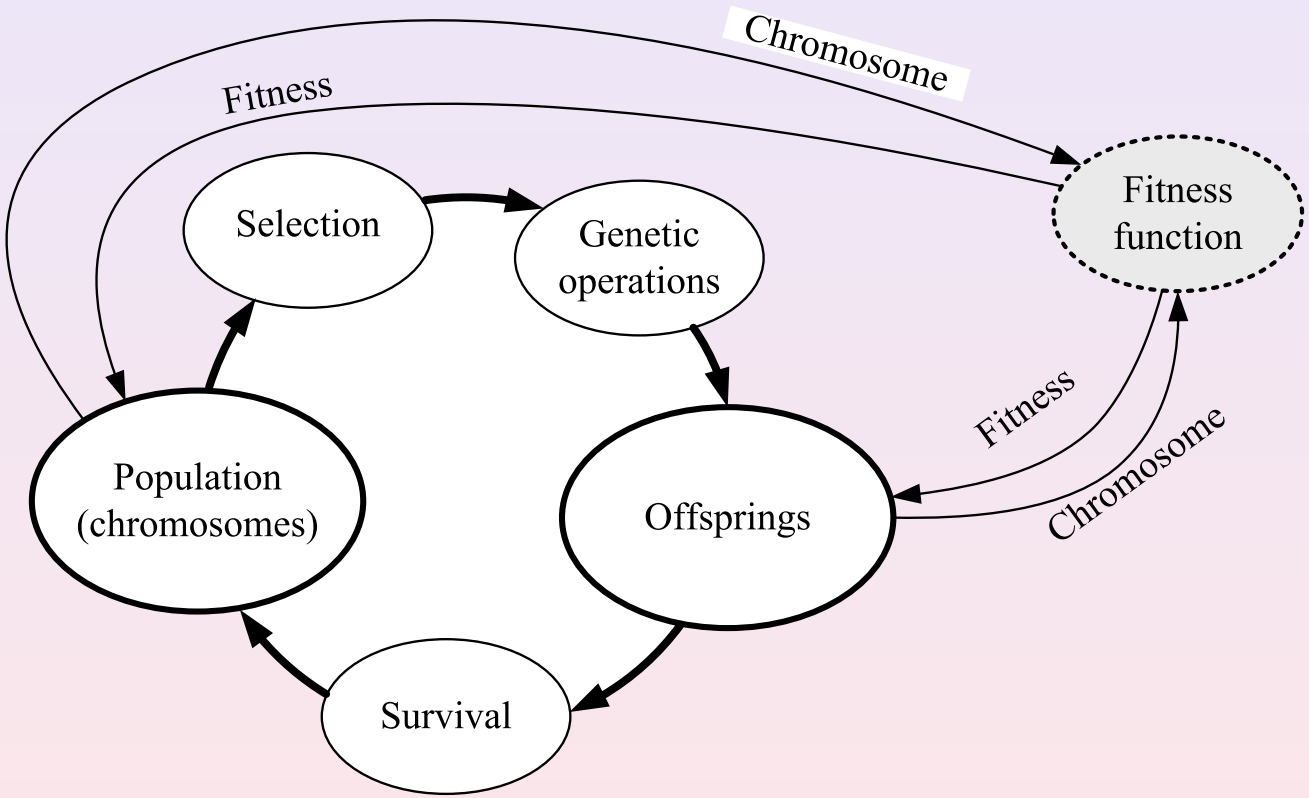


Figure: The genetic search process.

Typical tasks for ML algorithms

- Classify observations
  - Good/Bad, Healthy/Sick, Red/Green/Blue, A/B/C,...
- Estimate some value for observations (usually is called "regression" in statistics)
  - How good/bad is it?
  - How healthy is the patient?
  - How many points will I gain?
  - What is a reasonable price for this house?
  - etc.