An Introduction to Statistical Machine Learning - Introduction -

Samy Bengio

bengio@idiap.ch

Dalle Molle Institute for Perceptual Artificial Intelligence (IDIAP)

CP 592, rue du Simplon 4

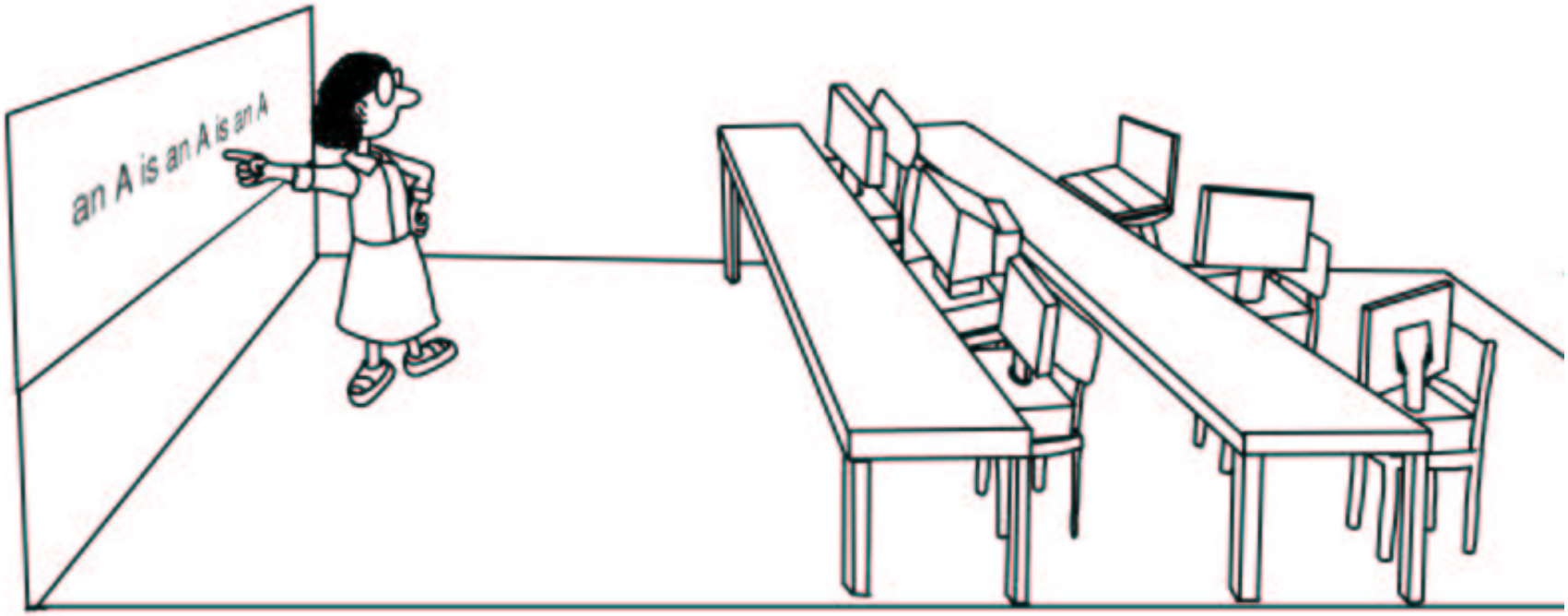
1920 Martigny, Switzerland

http://www.idiap.ch/~bengio

Introduction to Machine Learning

1. What is Machine Learning?
2. Why is it difficult?
3. Basic Principles
   (a) Occam's Razor
   (b) Learning as a Search Problem
4. Types of Problems
   (a) Regression
   (b) Classification
   (c) Density Estimation
5. Applications
6. Documentation

What is Machine Learning? (Graphical View)

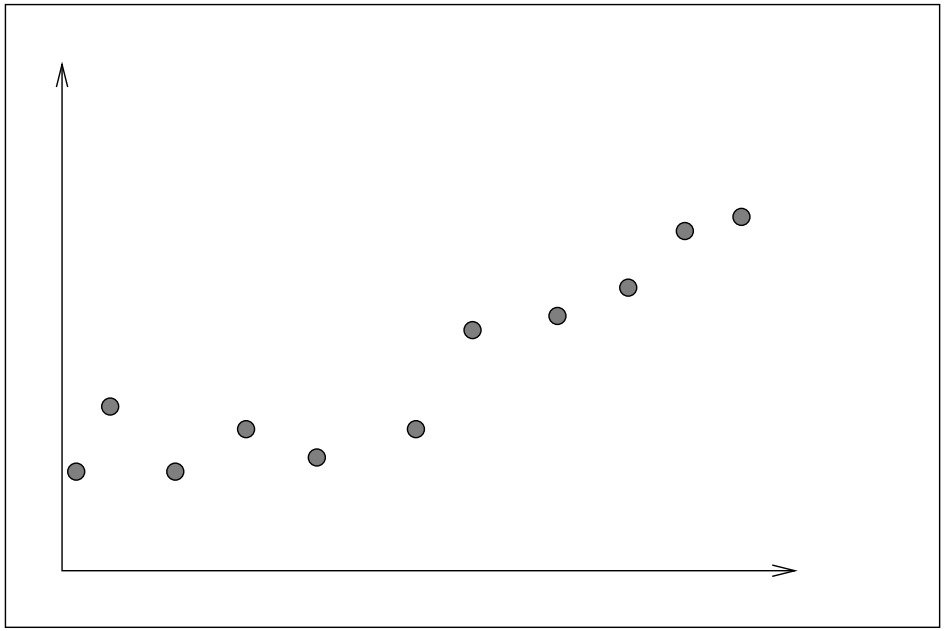


What is Machine Learning?

- Learning is an essential human property
- Learning means changing in order to be better (according to a given criterion) when a similar situation arrives
- Learning IS NOT learning by heart
- Any computer can learn by heart, the difficulty is to generalize a behavior to a novel situation

Why Learning is Difficult?

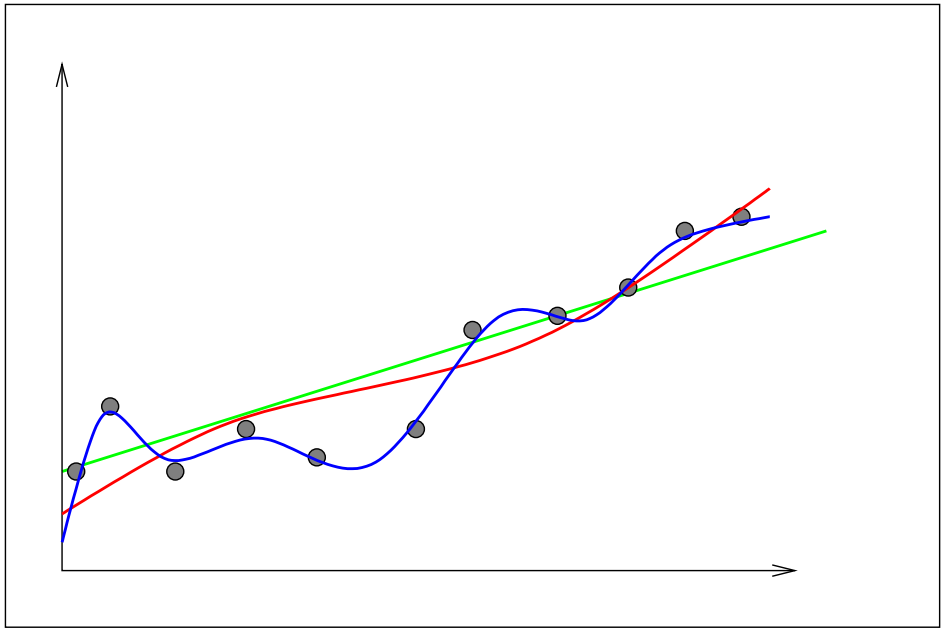
- Given a finite amount of training data, you have to derive a relation for an infinite domain
- In fact, there is an infinite number of such relations



- How should we draw the relation?

Why Learning is Difficult? (2)

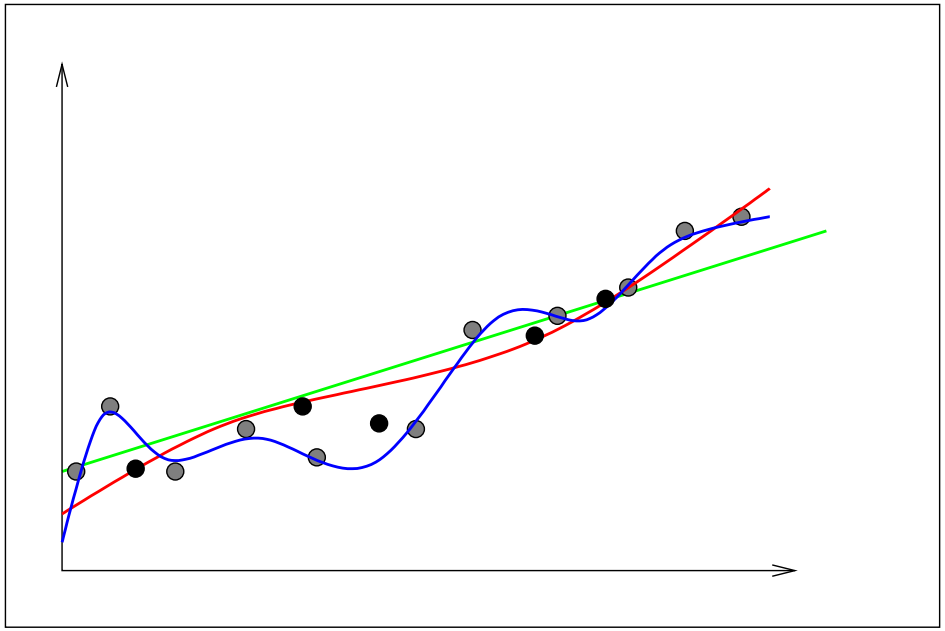
- Given a finite amount of training data, you have to derive a relation for an infinite domain
- In fact, there is an infinite number of such relations



- Which relation is the most appropriate?

Why Learning is Difficult? (3)

- Given a finite amount of training data, you have to derive a relation for an infinite domain
- In fact, there is an infinite number of such relations



- ... the hidden test points...

Occam's Razor's Principle

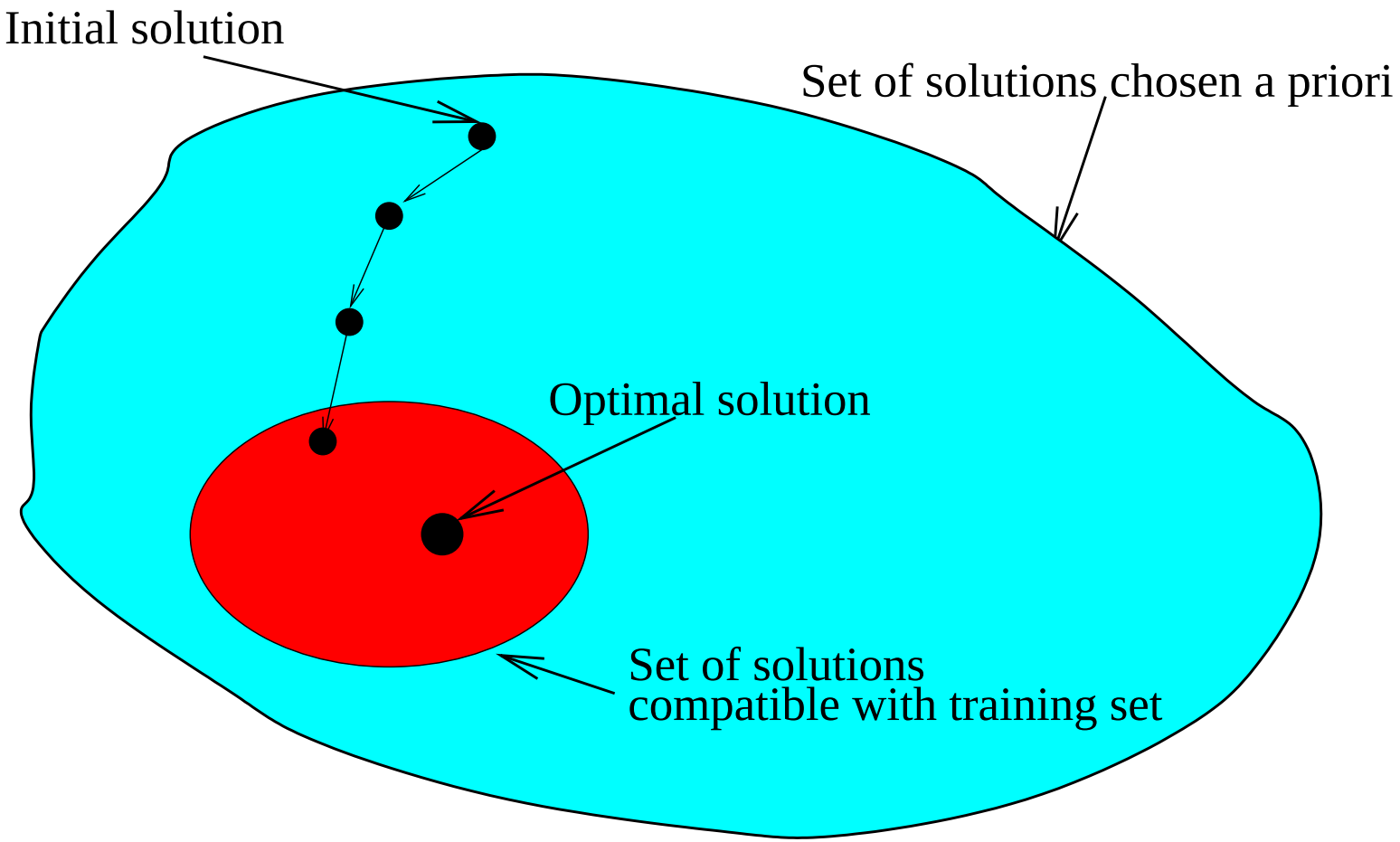
- William of Occam: Monk living in the 14th century
- Principle of Parcimony:

  One should not increase, beyond what is necessary, the number of entities required to explain anything

- When many solutions are available for a given problem, we should select the simplest one
- But what do we mean by simple?
- We will use prior knowledge of the problem to solve to define what is a simple solution

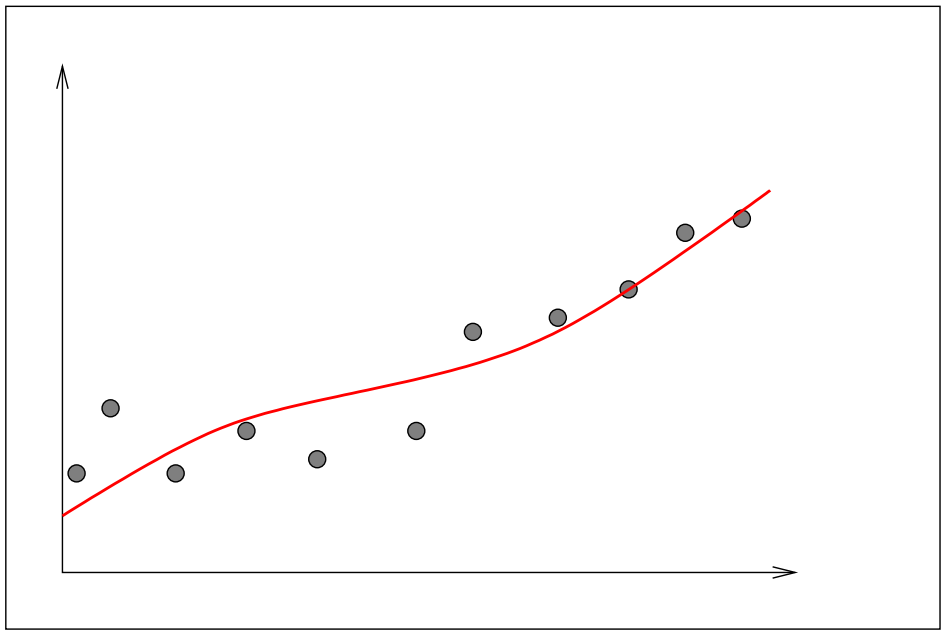
  Example of a prior: smoothness

Learning as a Search Problem



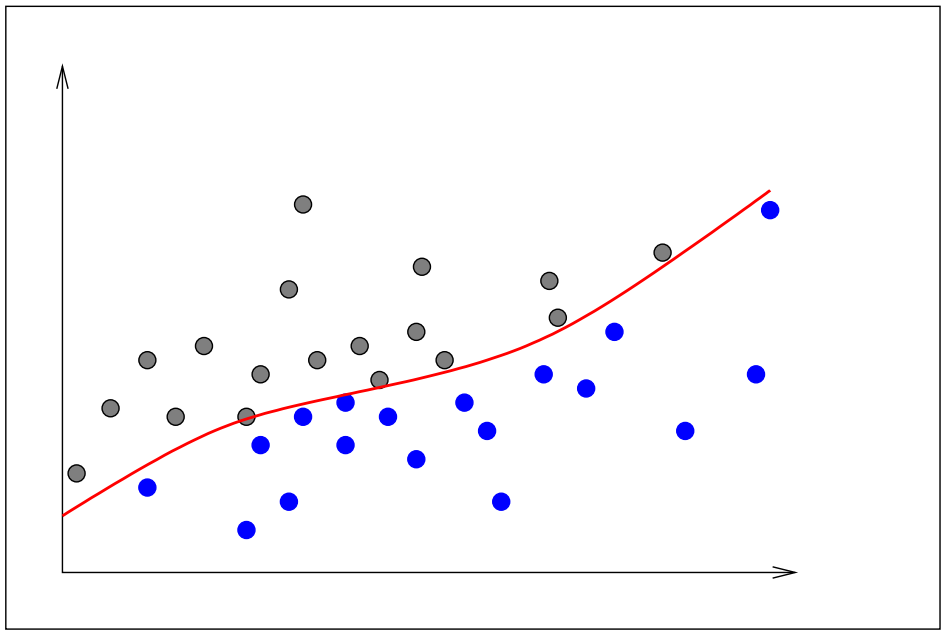
Types of Problem

- There are 3 kinds of problems:
  - regression



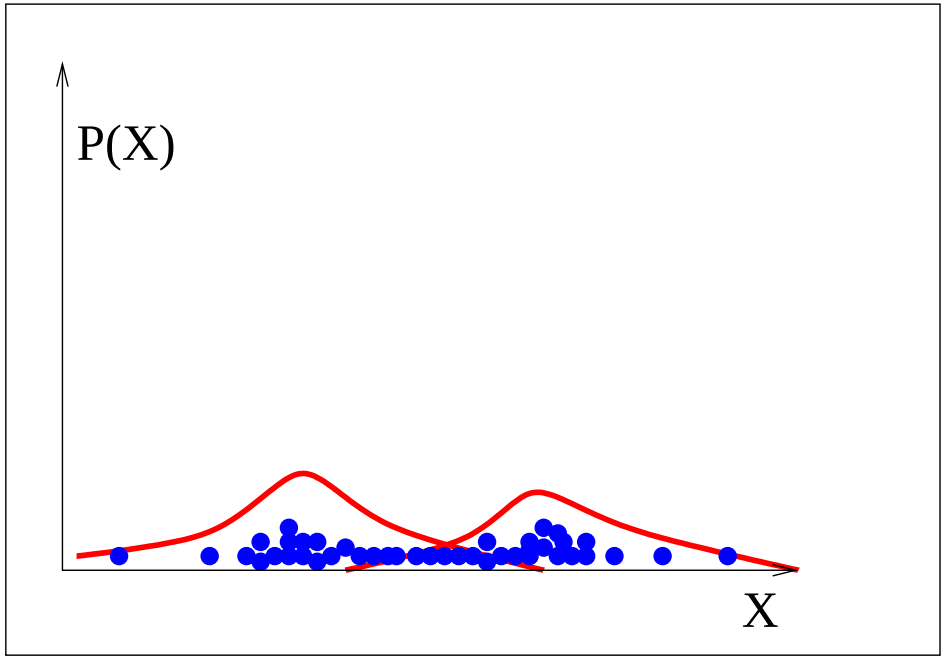
Types of Problem

- There are 3 kinds of problems:
  - regression, classification



Types of Problem

- There are 3 kinds of problems:
  - regression, classification, density estimation



Applications

- Vision Processing
  - Face detection/verification
  - Handwritten recognition
- Speech Processing
  - Phoneme/Word/Sentence/Person recognition
- Others
  - Finance: asset prediction, portfolio and risk management
  - Telecom: traffic prediction
  - Data mining: make use of huge datasets kept by large corporations
  - Games: Backgammon, go
  - Control: robots
- ... and plenty of others of course!

Documentation

- Machine learning library: www.Torch.ch
- Journals:
  - Journal of Machine Learning Research
  - Neural Computation
  - IEEE Transactions on Neural Networks
- Conferences:
  - NIPS: Neural Information Processing Systems
  - COLT: Computational Learning Theory
  - ICML: International Conference on Machine Learning
  - ICANN & ESANN: 2 European conferences
- Books:
  - Bishop, C. Neural Networks for Pattern Recognition, 1995.
  - Vapnik, V. The Nature of Statistical Learning Theory, 1995.

Documentation

- Search engines:
  - NIPS online: http://nips.djvuzone.org
  - NEC: http://citeseer.nj.nec.com/cs
- Other lecture notes: (some are in french...)
  - Bengio, Y.: http://www.iro.umontreal.ca/~bengioy/ift6266/
  - Kegl, B.: http://www.iro.umontreal.ca/~kegl/ift6266/
  - Jordan, M.: http://www.cs.berkeley.edu/~jordan/courses/294-fall98/