Neural Networks

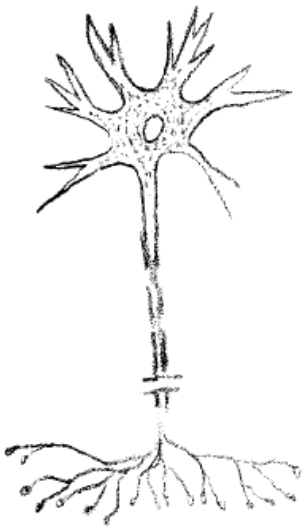
Marcus Dobrinkat Tino Ojala

Introduction

Biological Neural Networks

models used in

Theoretical Neuroscience



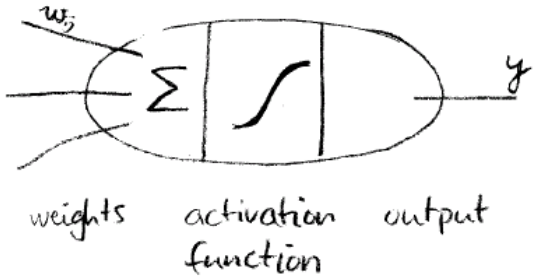
common features

parallel computations
distributed
non linear units
local processing
adaptation

Articficial Neural Networks

models used in

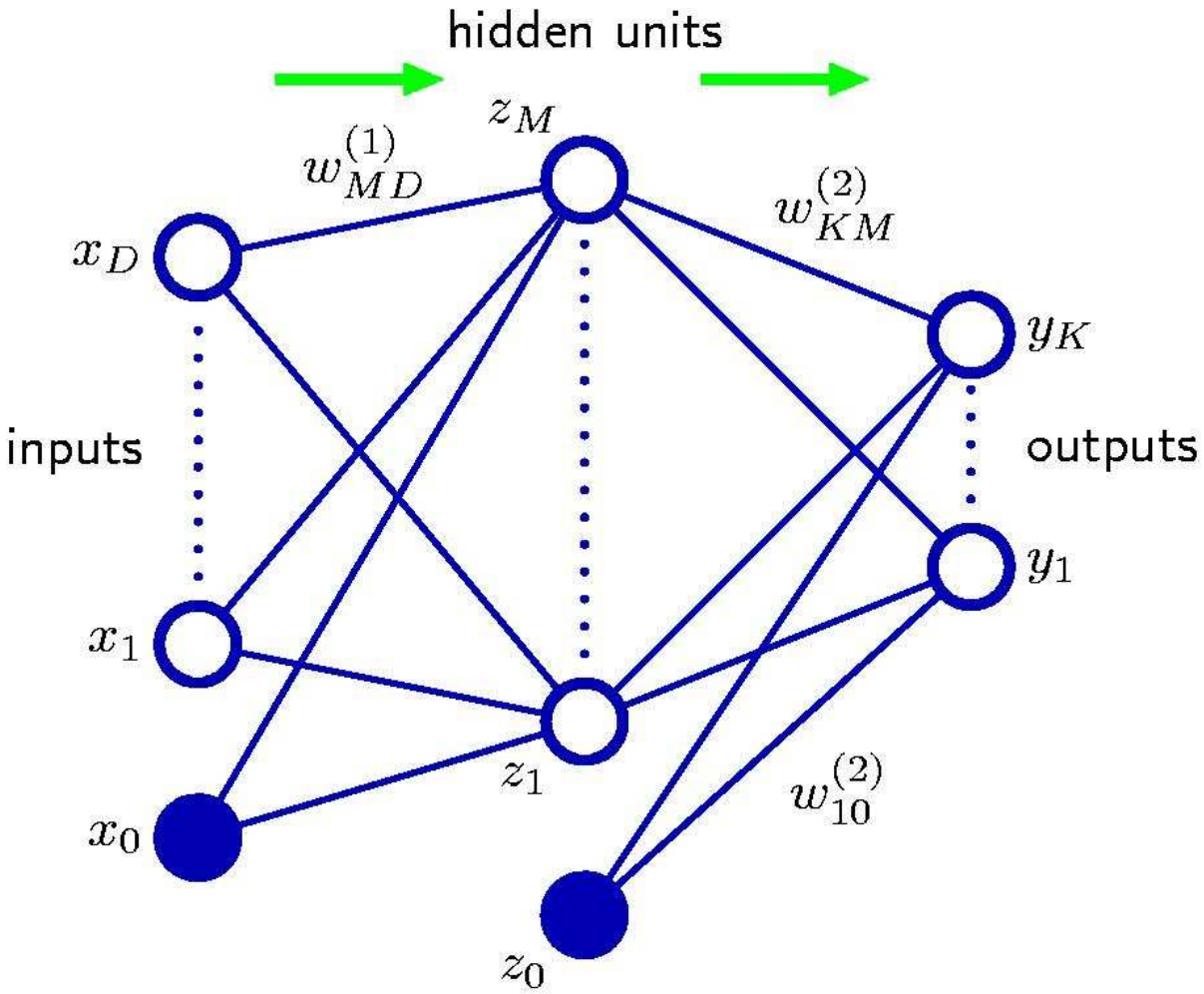
Function Approximation
Classification
Data Processing



Introduction

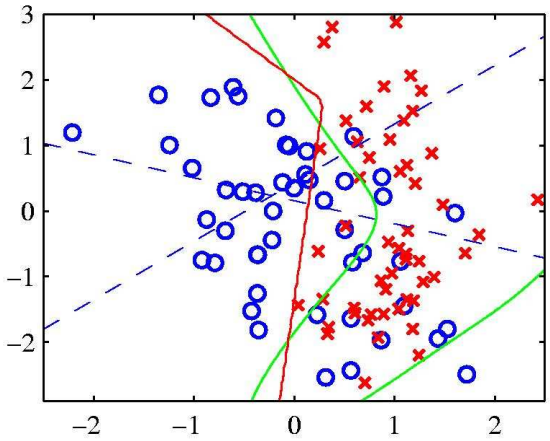
- Linearly separable problems
- XOR Problem
- Why a nonlinear function?
- 2 layers, 4 units

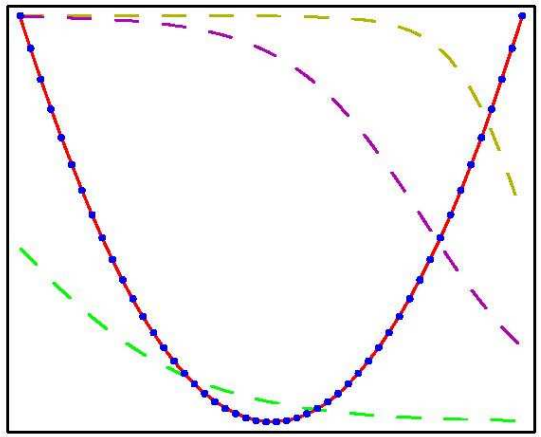
Multi Layer Networks



Network Training / Learning

- Classification vs. Regression





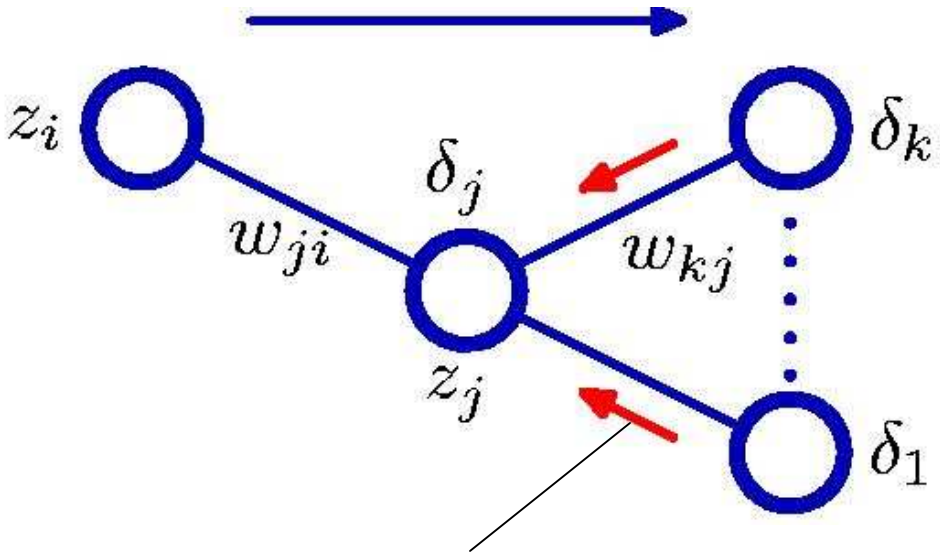
- Batch Learning vs Online Learning

Learning

- Online Learning
  - weights updated after every training sample
  - significantly faster than offline learning
  - better suited for large datasets
- Offline/Batch Learning
  - weights updated after one epoch

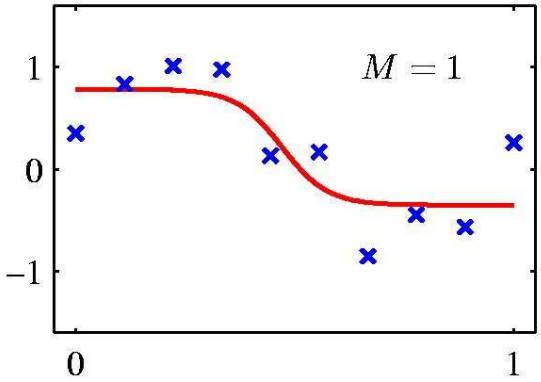
Backpropagation

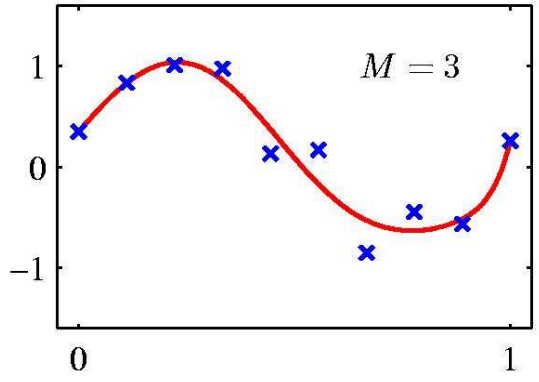
1. Forward Information Flow

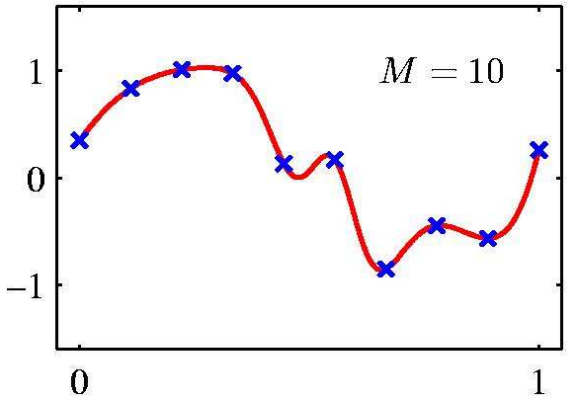


2. Errors Propagated Backwards

Regularization



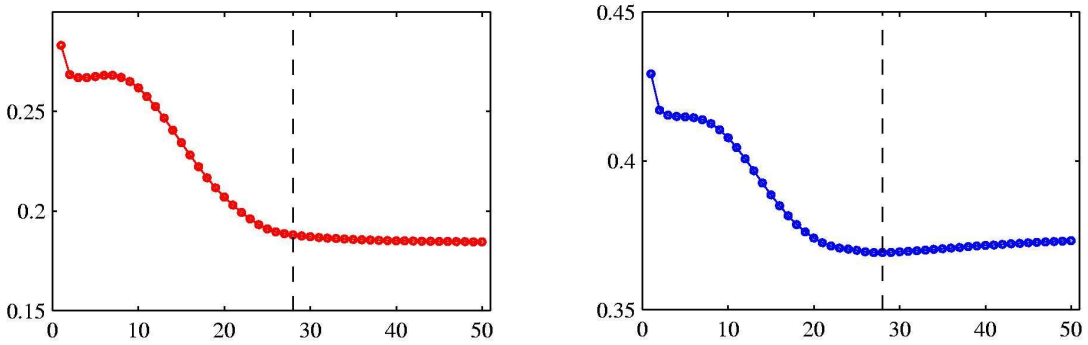




- suitable M or
- large M, but regularization

  e.g. early stopping

Early Stopping



Invariances

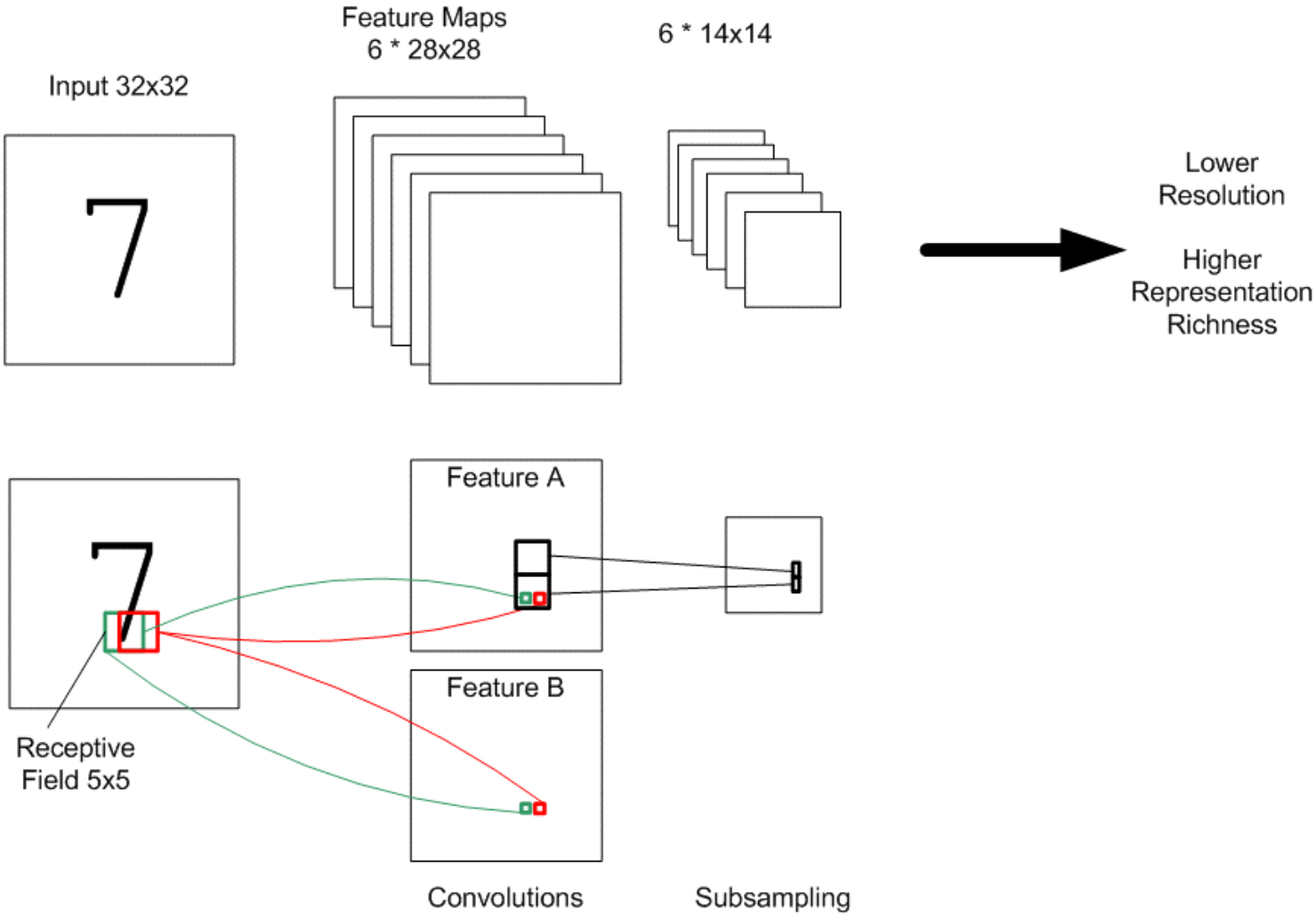
- Same output althought some variance in input.

How?

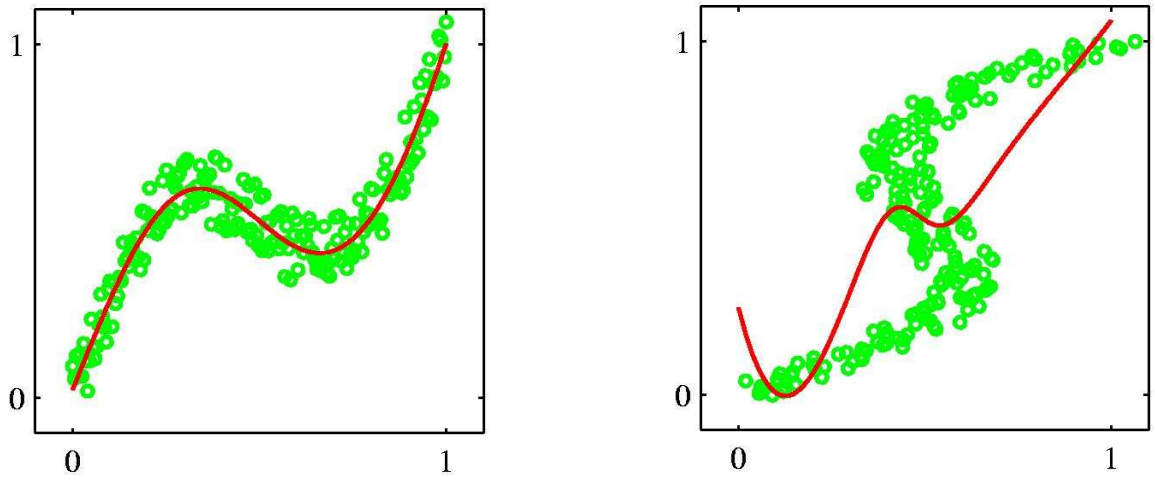
- all examples in training data
- varying copies
- preprocessing
- network itself

    -regularization term, receptive fields...

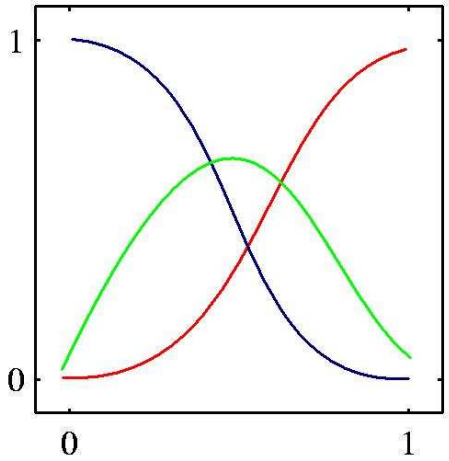
Convolutional Networks

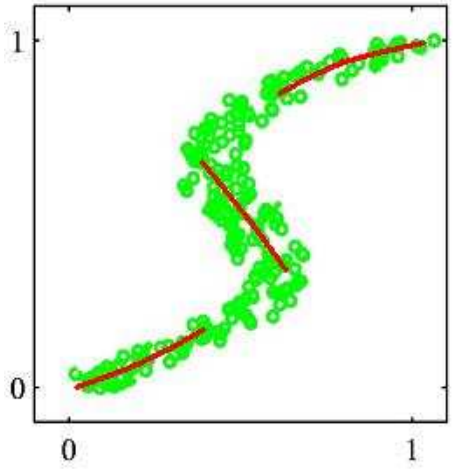


Mixture Desity Networks



- Gaussian assumption





output:

- 3 means
- 3 variances
- 3 mix. coeff

result

Questions?

References

- C. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
- LeNet5 Convolutional Networks http://yann.lecun.com/exdb/lenet/index.html
- Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "*Gradient-Based Learning Applied to Document Recognition*" *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, Nov. 1998.
- S. Haykin, *Neural Networks a Comprehensive Foundation* (2ed), Prentice-Hall, Englewood Cliff, NJ, 1999.
- S.Russell & P.Norwig, *Artificial Intelligence: a modern approach*, (2ed) Prentice Hall, 2003.
- Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*, Second Edition, John Wiley & Sons, New York, 2001.