

We call ourselves *Homo sapiens* -man the wise- because our intelligence is so important to us. For thousands of years, we have tried to understand how we think; that is, how a mere handful of matter can perceive, understand, predict, and manipulate a world far larger and more complicated than itself.

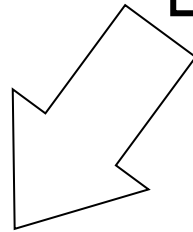
The field of artificial intelligence, or AI, goes further still: it attempts not just to understand but also to build intelligent entities.

<p>Thinking Humanly</p> <p>"The exciting new effort to make computers think ... machines with minds, in the full and literal sense." (Haugeland, 1985)</p> <p>"[The automation of] activities that associate with human thinking, activities such as decision-making, problem solving, learning ..." (Bellman, 1978)</p>	<p>Thinking Rationally</p> <p>"The study of mental faculties through the use of computational models" (Charniak and McDermontt, 1985)</p> <p>"The study of the computations that make it possible to perceive, reason and act" (Winston, 1992)</p>
<p>Acting Humanly</p> <p>"The art of creating machines that perform functions that require intelligence when performed by people" (Kurzweil, 1990)</p> <p>"The study of how to makes computers do things at which, at the moment, people are better" (Rich and Knight, 1991)</p>	<p>Acting Rationally</p> <p>"Computation Intelligence is the study of the design of intelligent agents" (Poole et al., 1998)</p> <p>"AI ... is concerned with intelligent behavior in artifacts" (Nilsson, 1998)</p>

PREREQUISITES

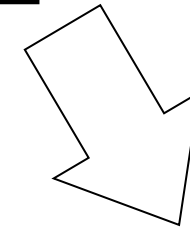
LINEAR ALGEBRA

1. Dot Product
2. Singular Value Decomposition
3. The Moore-Penrose Pseudoinverse



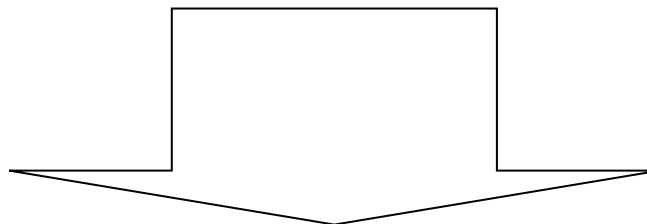
PROBABILITY THEORY

1. Conditional Probability
2. Expectation, Variance and Covariance
3. Bayes' Rule
4. Common Probability Distribution



NUMERICAL COMPUTATION

1. Gradient-Based Optimization
2. Linear Least Squares



DENIS ENĂCHESCU

University of Bucharest

SUPERVISED LEARNING of NEURAL NETWORKS

CONTENTS

1. THE NATURE OF MACHINE LEARNING

1.1. Basic Definitions and Key Concepts

1.2. Short History

1.3. General Model of Learning from Examples

1.3.1. The Theory of Inductive Inference

1.3.2. How to Analyze the Learning

1.3.3. Validity Conditions for the ERM Principle

1.3.4. The Bayesian Model of Learning

1.4. The Bias-Variance Tradeoff

1.5. Regularization Methods

1.5.1. Tuning the Hypotheses Class

1.5.2. Selection of the Models

1.6. Evaluation of the Learning Performances

1.6.1. A Posteriori Empirical Evaluation

1.6.2. Practical Selection of the Model

1.6.3. Estimation of the Hypothesis Real Risk

1.6.4. Algorithms Tuning by a Validation Set

1.7. Comparison of the Learning Methods

- 1.7.1. Comparison of two Hypotheses Produced by the same Algorithm on two Different Test Samples
- 1.7.2. Comparison of two Algorithms on Different Test Sets
- 1.7.3. Comparison of two Algorithms on the Same Test Sets

2. SUPERVISED LEARNING OF A PERCEPTRON

2.1. Error Correction Rules

- 2.1.1. Perceptron Learning Rule
- 2.1.2. Generalizations of the Perceptron Learning Rule
- 2.1.3. The Perceptron Criterion Function
- 2.1.4. May's Learning Rule
- 2.1.5. Widrow-Hoff (α -LMS) Learning Rule

2.2. Other Gradient-Descent-Based Learning Rules

- 2.2.1. μ -LMS Learning Rule
- 2.2.2. The μ -LMS as a Stochastic Processorrelation Learning Rule

2.3. The Delta Rules

2.4. Adaptive Ho-Kashyap (AHK) Learning Rules

2.5. Other Criterion Function

2.6. Summary

3. LEARNING RULE FOR MULTILAYER FEEDFORWARD NEURAL NETWORKS

3.1. Error Backpropagation Learning Rule

3.2. Global-Descent-Based Error Backpropagation

3.3. Backprop Enhancements and Variation

3.3.1. Weights Initialization

3.3.2. Learning Rate

3.3.3. Momentum

3.3.4. Activation Function

3.3.5. Weight Decay, Weight Elimination and Perceptron Elimination

3.3.6. Cross-Validation

3.3.7. Criterion Functions

4. COMPUTATION CAPABILITIES OF NEURAL NETWORKS

4.1. Bounds on the Number of Functions Realizable by a Feedforward Networks of LTGs

4.2. Necessary Lower Bounds on the Size of Neural Networks

- 4.2.1. Two Layer Feedforward Networks
- 4.2.2. Three Layer Feedforward Networks
- 4.2.3. Generally Interconnected Networks with No Feedback

4.3. Approximation Capabilities of Feedforward Neural Networks

- 4.3.1. Kolmogorov's Theorem
- 4.3.2. Single-Hidden-Layer Neural Networks Are Universal Approximators
- 4.3.3. Single-Hidden-Layer Neural Networks Are Universal Classifiers

BIBLIOGRAPHY

S.J. Russel, P. Norvig, *Artificial Intelligence. A Modern Approach*, Third Edition, Pearson Education Ltd, 2014

T. Fine, *Feedforward Neural Network Methodology*, Springer, 1999

S. Haykin, *Neural Network and Learning Machines*, Third Edition, Pearson Education Ltd, 2009

M.T. Hagan, H.B. Demuth, M. Beale, O.de Jesu, *Neural Network Design*

Follow up

- Master studies
- PhD studies
- Employment in big companies
 - <https://www.humanbrainproject.eu/discover/the-project/research-areas>
 - <http://www.darpa.mil/our-research>
 - <https://blogs.microsoft.com/next/2016/07/07/project-malmo-lets-researchers-use-minecraft-ai-research-makes-public-debut/#sm.0001re0afyfxpdy3qfc2862sk6qij>
 - <https://deepmind.com/>
 - <https://www.palantir.com/> Palantir is one of the most prominent companies in the relatively new field of IA and in strong expansion. The company was founded in 2004 by Peter Thiel (PayPal co-founder, board of directors Facebook, one of the main VIP in Silicon Valley).

- <http://www.aria-romania.org>
 - BITDEFENDER (machine learning),
 - Avira (machine learning),
 - TeamNet (machine learning, multi-agents systems, robotics, etc),
 - Oracle (natural language processing).
 - Eau de Web (semantic web and linked open data).
 - AQUAsoft (ambient intelligence).
- - Recognos (semantic web), Cluj.