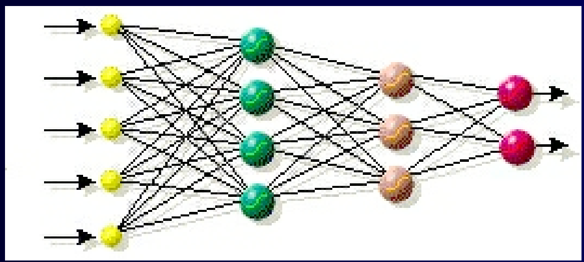
Capabilities, limitations and fascinating applications of Artificial Neural Networks



SURVEY OF LECTURE

- Definition of concepts :neuron, neural network, training, learning rules, activation function
- Feedforward neural network
- Multilayer perceptron
- Learning, generalization, early stopping
- Training set, test set
- Overtraining
- Comparison, digital computer, artificial neural network
- Comparison, artificial neural networks, biologic brain
- History of neural networks
- Application fields of neural networks
- Overview of case studies
- Practical advice for successful application
- Internet references
- Prospects of commercial use

Fascinating applications, capabilities and limitations of Artificial neural networks : 6 objectives

- artificial neural network not magic, but design based on solid mathematical methods
- difference : neural networks versus computers limitations of artificial neural networks versus the human brain
- neural networks better than computer for proces-sing of sensorial data such as signalprocessing, image processing, pattern recognition, robotcontrol, non-linear modeling and prediction

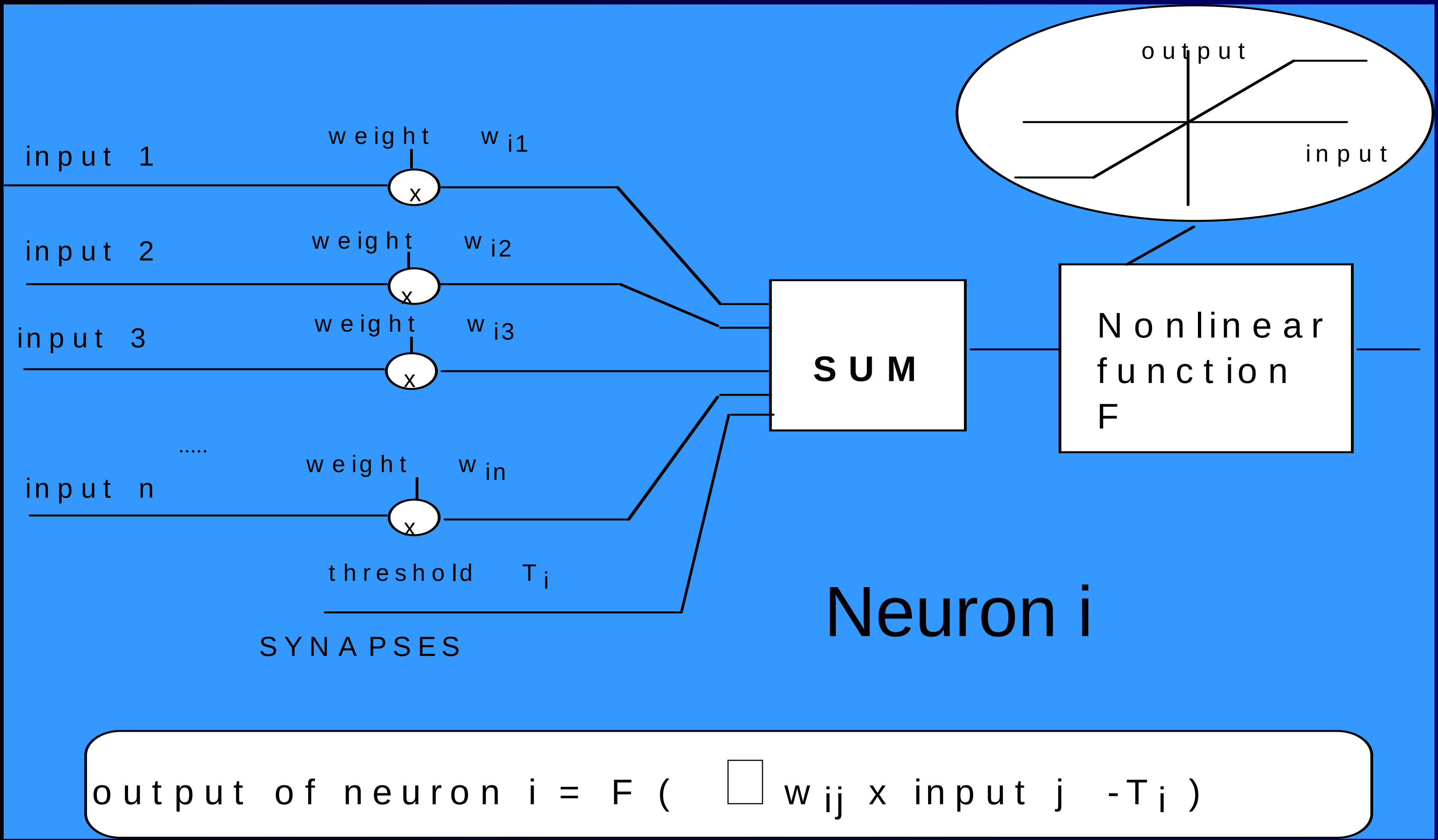
6 objectives

- survey of attractive applications of artificial neural networks.
- practical approach for using artificial neural netwerks in various technical, organizatorial and economic applications.
- prospects for use of artificial neural networks in products

Ambition : to understand the mathematical equations, and the role of the various parameters

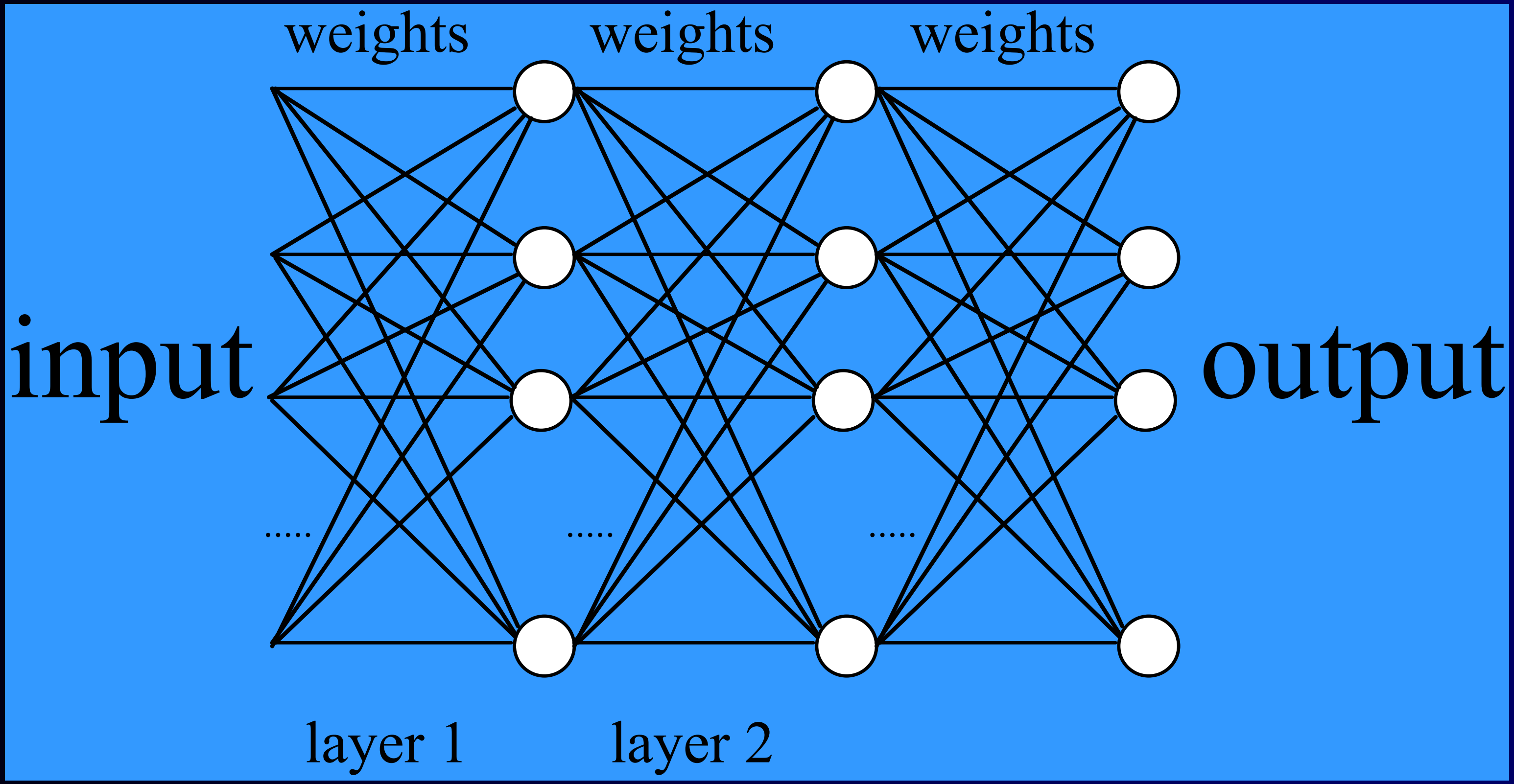
What is a neuron ?

neuron makes a weighted sum of inputs and applies a non-linear activation function.



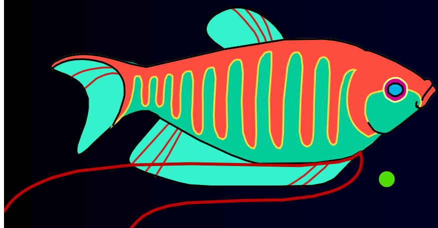
What is a neural network ?

“artificial” neural network= mathematical model of network with neurons.
≠ biologic neural networks (much more complicated)

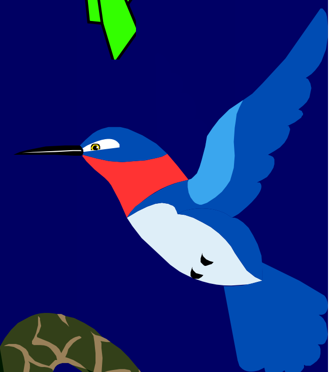


Universal approximation property

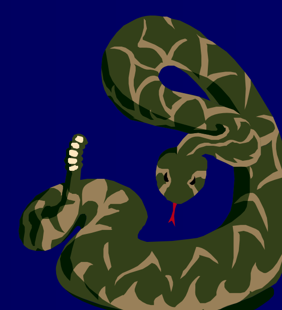
Learning = adapting weights with examples


- weights adapted during learning or training.
- learning rule adaptation of the weights according to the examples.
- a neural network learns from examples

- eg. children classify animals from living examples and photographs
- neural networks obtain their information during the learning process and store the information in the weights.

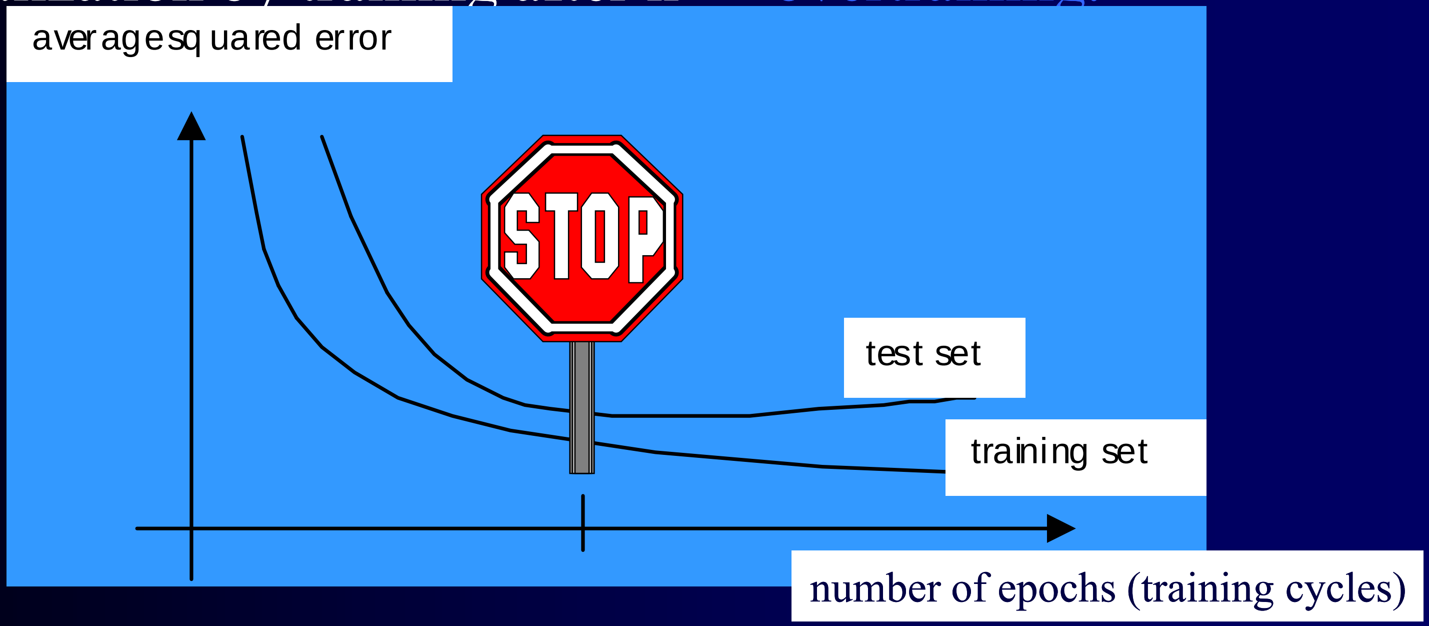
- But, a neural network can learn something unexpected

Learning and testing

- adapting the weights by Back propagation of the error : one applies one by one the fraud examples to the inputs of the neural network and checks if the corresponding output is high.
  If so then no adaption,
  if not, then adaption of weights according to the learning rule. Keep applying the examples until sufficiently accurate decisions are made by the neural network (stop rule) : often many rounds or epochs.
- use of trained network: apply during the night the operations of the previous day to find the few fraud cases out of millions of cards --> no legal proof, but effective
  neural networks are implicitly able to generalize , i.e. the neural network can retrieve similar fraud cases.

generalization property

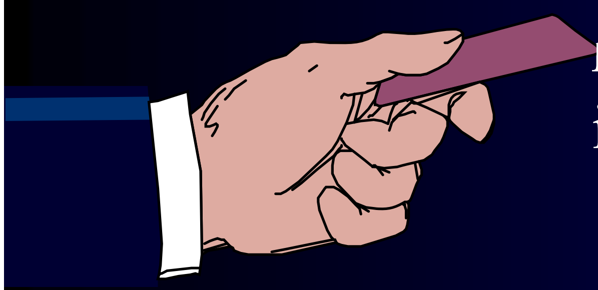
- partition the collection of credit card data records into 2 sets
- learning set = training set for adapting the weights during learning -->decrease in error
- test set typically first decrease, then slight increase: worse generalization by training after n -->overtraining.



Stop when the error for the test set increases i.e. as long as the neural network generalizes well.

Example of an application of neural networks

- detecting fraud with credit cards.
  objective : detect fraud as soon as possible in a dataset of millions of cards.
- expertsystems = collection of rules that describe fraudulent behaviour explicitly--> problems
- alternative approach : neural networks : large collection of frauds for training a forward neural network with 3 layers i.e. apply actions of creditcard users at the input of the first layer of neurons. When a certain neuron in the output layer is high, then fraud of a certain type is detected.



Conclusion and warning from example

- misconception of users: use test set also during training. --> no correct prediction of the crucial generalization property of the neural network
- use of neural networks : modeling and computation for every function and many technical and non-technical systems -->the neural network can approximate every continuous mapping between inputs and outputs (universal approximation property) --> practically : neural networks are interesting whenever examples are abundant, and the problem cannot be captured in simple rules.

digital computer vs neural network

- working principle symbols, “1” or “0” /program Von Neumann principle / mathematical logic and Boolean algebra/ programs software / algorithms, languages, compilers, design methodologies
- parallellisation difficult : sequential processing of data
- useless without software
- rigid : modify one bit, disaster
- conclusion: important differences

-->new paradigm for information processing

- working principle patterns / learn a nonlinear map/ mathematics of nonlinear functions or dynamical systems/ need for design methodologies
- parallellisation easy parallel by definition cfr brain
- useless without training
- choice of learning rule and examples crucial
- robust against inaccuracies in data, defect neurons and error-correcting capability ->collective behavior cfr brain

neural networks vs human brains

- low complexity : electronic VLSI chip : < few thousand neurons on 1 chip / simulations on computers : few 100.000 neurons
- high processing speed : 30 to 200 million basic operations per sec on a computer or chip
- energetic efficiency : best computers now consume 10**-6 Joule per operation and per sec
- conclusion : methodology for design and use of neural networks ≠ biologic neural networks

- high complexity : human brain 100.000.000.000 neurons --> gap cannot be bridged in a few decennia
- low processing speed : reaction time of biologic neural networks : 1 to 2 millisec.
- energetic efficiency : biologic neural network much better. 10**-16 Joule per operation and per sec
- conclusion : modesty with respect to the human brain

neural networks vs human brains

- analogy with biologic neural networks is too weak to convince engineers and computer scientists about correctness.
- correctness follows from mathematical analysis of non-linear functions or dynamical systems and computer simulations.

History of Neural Networks

- 1942 Mc Culloch and Pitts : mathematical models for neurons
- 1949 psychologist Hebb first learning rule--> memorize by adapting weights
- 1958 Rosenblatt : book on perceptrons : a machine capable to classify information by adapting weights
- 1960-62 Widrow and Hoff : adaline and LMS learning rule
- 1969 Minsky and Papert prove limitations of perceptron
- 13 years of hibernation!! but some stubborn researchers Grossberg(US), Amari and Fukushima(Japan), Kohonen(Finland) and Taylor(UK)
- 1982 Kohonen describes his self-organizing map
- 1986 Rumelhart rediscovers backpropagation
- ≥ 1987 much research on neural networks, new journals, conferences, applications, products, industrial initiatives, startup companies

Fascinating applications and limitations of neural networks

- Neural networks -->cognitive tasks : processing of several sensorial data, vision, image and speech processing, robotics, control of objects and automation.
- Digital computers -->rigid tasks : electronic spreadsheets, accountancy, simulation, electronic mail, text processing
- complementary application fields: combined use.
- many convincing applications of neural networks -->abundant literature (hundreds of books, dozen of journals, and more than 10 conferences per year). For novice practical guidelines without much mathematics and close to application field. For expert many journals and conference papers

survey of application categories

- expertsystems with neural networks fraud detection with credit cards, fraud detection with mobilophony, selection of materials in certain corrosive environments and medical diagnosis.
- pattern recognition : speech, speech-controlled computers, en telephony, recognition of characters and numbers, faces and images: recognition of handwriting, addresses on envelopes,searching criminal faces in a database, recognition of car license plates, …

  special chips e.g. cellullar neural networks only connection to neighboring neurons in a grid.Every neuron processes one pixel and has one ligth-sensitive diode -->future prospect of artificial eye
- optimization of quality and product and control of mechanical, chemical and biochemical processes the non-linearity of the neural network provides improvements w.r.t. traditional linear controllers for inherently non-linear systems like the double inverse pendulum (chaotic system).
- prediction not "magic" : exchange rates, portfolio -->improvements from 12.3 % to 18 % per year, prediction of electricity consumption crucial in electrical energy sector, no storage of electrical energy: production = consumption

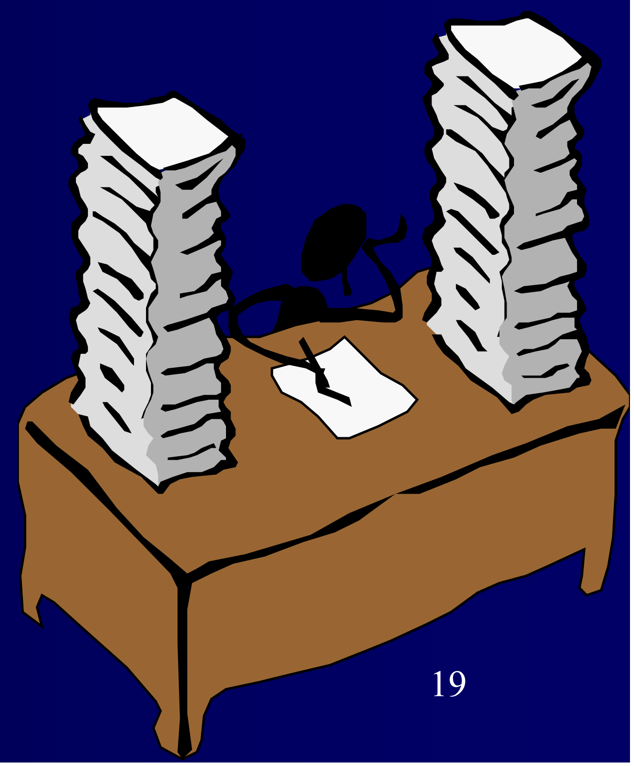
autonomous vehicle control with a neural network

(ALVINN project).

- goal: keep the vehicle without driver on the road car equipped with videorecorder with 30 x 32 pixels and a laserlocalizer that measures the distance between the car and the environment in 8 x 32 points.
- the architecture of the neural network 30 x 32 + 8 x 32 = 1216 measurements of inputs and outputs. hidden layer of 29 neurons and an output layer of 45 neurons. steering direction of the car : middle neuron highest --> straight forward. Most right neuron highest, maximal turn right and analogously for left learning phase recording 1200 combinations of scenes, light and distortions with human driver. neural network trained and tested in about half an hour computing time with backpropagation --> quality of driving up to 90 km/h comparable to the best navigation systems
- major advantage of neural networks fast development time Navigation systems require a development time of several months for design and test of vision-software, parameter-adaptations, and program- debugging, short development time because the neural network can capture the essential features of a problem without explicit formulation.

Datamining with neural networks

- Data definition and collection important
- Choice of variables
- Incomplete data better than incorrect data
- Negative as well as positive examples needed
- Coding of the outputs important



Case studies of successful applications

Stimulation Initiative for European Neural Applications Esprit Project 9811

- Benelux
- Prediction of Yarn Properties in Chemical Process Technology
- Current Prediction for Shipping Guidance in IJmuiden
- Recognition of Exploitable Oil and Gas Wells
- Modelling Market Dynamics in Food-, Durables- and Financial Markets
- Prediction of Newspaper Sales
- Production Planning for Client Specific Transformers
- Qualification of Shock-Tuning for Automobiles
- Diagnosis of Spot Welds
- Automatic Handwriting Recognition
- Automatic Sorting of Pot Plants
- Spain/Portugal
- Fraud detection in credit card transactions
- Drinking Water Supply Management
- On-line Quality Modelling in Polymer Production
- Neural OCR Processing of Employment Demands
- Neural OCR Personnel Information Processing
- Neural OCR Processing of Sales Orders
- Neural OCR Processing of Social Security Forms

Case studies of successful applications(cont.)

- Germany/Austria
- Predicting Sales of Articles in Supermarket
- Automatic Quality Control System for Tile-making Works
- Quality Assurance by "listening"
- Optimizing Facilities for Polymerization
- Quality Assurance and Increased Efficiency in Medical Projects
- Classification of Defects in Pipelines
- Computer Assisted Prediciton of Lymphnode-Metastasis in Gastric Cancer
- Alarm Identification
- Facilities for Material-Specific Sorting and Selection
- Optimized Dryer-Regulation
- Evaluating the Reaction State of Penicillin-Fermenters
- Substitution of Analysers in Distillation Columns
- Optical Positioning in Industrial Production
- Short-Term Load Forecast for German Power Utility
- Monitoring of Water Dam
- Access Control Using Automated Face Recognition
- Control of Tempering Furnaces
- France/Italy
- Helicopter Flight Data Analysis
- Neural Forecaster for On-line Load Profile Correction
- UK/Scandinavia
- For more than 30 UK case studies see DTI's NeuroComputing Web

successful applications at KULeuven/ICNN

- modelling and prediction of gas and electricity consumption in Belgium
- diagnosis of corrosion and support of metal selection
- modelling and control of chemical processes
- modelling and control of fermentation processes
- temperature compensation of machines
- control of robots
- control of chaotic systems
- Dutch speech recognition
- design of analog neural chips for image processing
- diagnosis of ovarian cancer
- fraud detection/ customer profiling

Practical advices for successful application

- creation of training and test set of examples : requires 90 % of time and effort. Bad examples-->bad neural networks / analyse data (correlations, trends, cycles) eliminate outliers, trend elimination, noise reduction, appropriate scaling, Fourier transform, and eliminating old data / how many examples? enough in order to have a representative set / rule of thumb : # examples in learning set = 5 X # weights in neural network / # examples in test set =#examples in learning set /2 / separation of learning set and test set arbitrary
- learning and testing: learning as long as the error for the test set decreases. If the neural network does not learn well, then adapt the network architecture or the step size. aim of learning: network should be large enough to learn and small enough to generalize evaluate the network afterwards because the neural network can learn something other than expected

Practical advices for successful application

- type of network : 3 layer feed-forward neural network /non-linearity: smooth transition from negative saturation (-1) for strongly negative input to positive saturation (+1) for strongly positive input. Between -1 and +1 active region neuron not yet committed and more sensitive to adaptations during training
- learning rule : error back propagation weights are adapted in the direction of the steepest descent of the error function i.e.weights are adapted such that the prediction errors of the neural network decrease stepsize choice of the user: if too small, cautious but small steps--> sometimes hundreds of thousands of cycles of all examples in the learning set are required. if too large, faster learning, but danger to shoot over the good choices
- size of the network : rule of thumb: # neurons of the first layer = #inputs/ #neurons in the third layer =#classes/ # neurons in middle layer not too small: no bottleneck/ too many neurons -->excessive computation time. e.g. 10.000 weights between two layers each with 100 neurons, adaptation of the weights with a learning set of 100 to 1000 examples a few seconds on a computer with 10**7 mult./s. and a few thousand training cycles --> few hours of computer time / too large a network--> overtraining : network has too many degrees of freedom too small a network : bad generalization.

Internet : frequently asked questions

World Wide Web http://www.faqs.org/faqs/ai-faq/neural-nets/part1/

- 1. What is this newsgroup for? How shall it be used?
- 2. What is a neural network (NN)?
- 3. What can you do with a Neural Network and what not?
- 4. Who is concerned with NNetworks?
- 5. What does 'backprop' mean? What is 'overfitting'?
- 6. Why use a bias input? Why activation functions?
- 7. How many hidden units should I use?
- 8. How many learning methods for NNs exist? Which?
- 9. What about Genetic Algorithms?
- 10. What about Fuzzy Logic?
- 11.Relation NN / statistical methods?
- 12. Good introductory literature about Neural Networks?
- 13. Any journals and magazines about Neural Networks?
- 14. The most important conferences concerned with Neural Networks?
- 15. Neural Network Associations?
- 16. Other sources of info about NNs?
- 17. Freely available software packages for NN simulation?
- 18. Commercial software packages for NN simulation?
- 19. Neural Network hardware?
- 20. Database for experiment with NN?

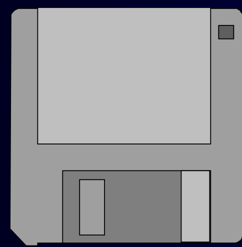
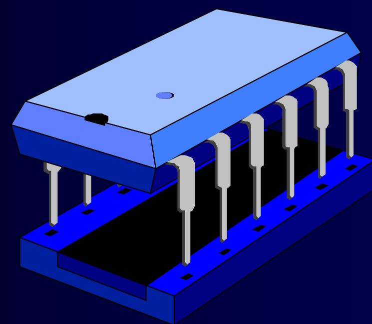
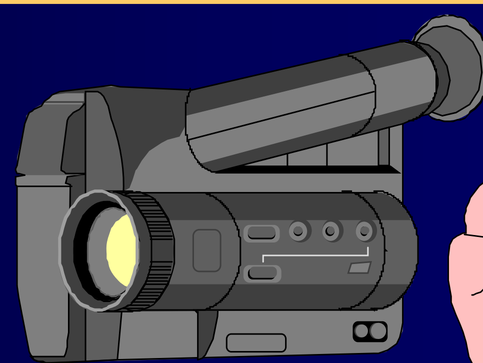
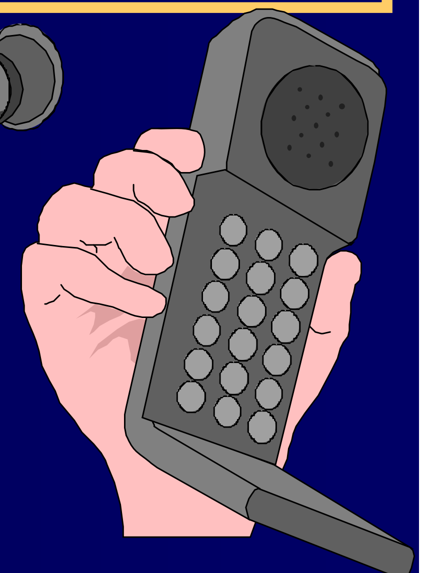
Subject: # Help! My NN won't learn! What should I do?

advice for inexperienced users. Experts may try more daring methods.

If you are using a multilayer perceptron (MLP):

Check data for outliers. Transform variables or delete bad cases
Standardize quantitative inputs see "Should I standardize the input variables?"
Encode categorical inputs see "How should categories be encoded?"
Make sure you have more training cases than the total number of input units.
at least 10 times as many training cases as input units.
Use a bias term ("threshold") in every hidden and output unit.
Use a tanh (hyperbolic tangent) activation function for the hidden units.
If possible, use conventional numerical optimization techniques see "What are conjugate gradients, Levenberg-Marquardt, etc.?"
If you have to use standard backprop, you must set the learning rate by trial and error. Experiment with different learning rates.
if the error increases during training, try lower learning rates.
When the network has hidden units, the results of training may depend critically on the random initial weights.

Prospects for commercial exploitation

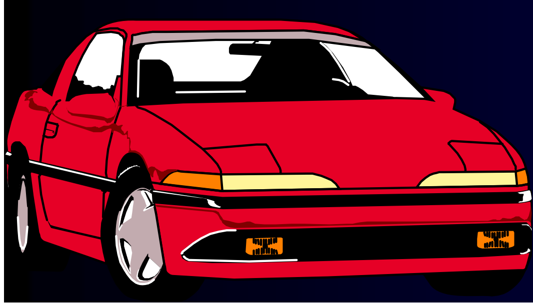






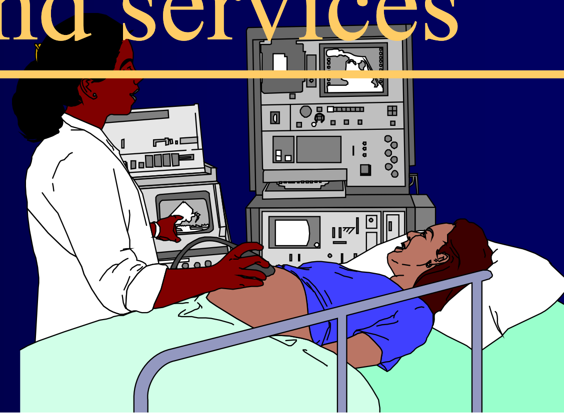
Traditional paradigm : Computer or chips + software

= Products and services

Advanced data processing and learning systems :
Computer or chips + examples

= Better Products and services




Conclusions

- Neural networks are realistic alternatives for information problems (in stead of tedious software development)
- not magic, but design is based on solid mathematical methods
- neural networks are interesting whenever examples are abundant, and the problem cannot be captured in simple rules.
- superior for cognitive tasks and processing of sensorial data such as vision, image- and speech recognition, control, robotics, expert systems.
- correct operation biologic analogy not convincing but mathematical analysis and computer simulations needed.
- technical neural networks ridiculously small w.r.t. brains good suggestions from biology
- fascinating developments with NN possible : specificities of the user voice-controlled apparatus, and pen-based computing