Neural Networks and Evolving Connectionist Systems. Applications for Classification and Prediction. An Introduction to the NeuCom Data Mining Environment

Prof. Nik Kasabov
Knowledge Engineering and Discovery Research Institute and
School of Computer and Information Sciences
Auckland University of Technology
nkasabov@aut.ac.nz
http://www.kedri.aut.ac.nz

Content

1. Data mining and knowledge engineering.
2. Neural networks: inspiration from the brain and general principles
3. Evolving connectionist and hybrid systems
4. Applications: Classification, prediction
5. Introduction and demonstrations of the NeuCom data mining environment.
6. References:

N.Kasabov (ed) Springer Handbook of Bio-/Neuroinformatics, 2014 (available free as a .pdf from the AUT Library as an eBook and alos as hard copy)

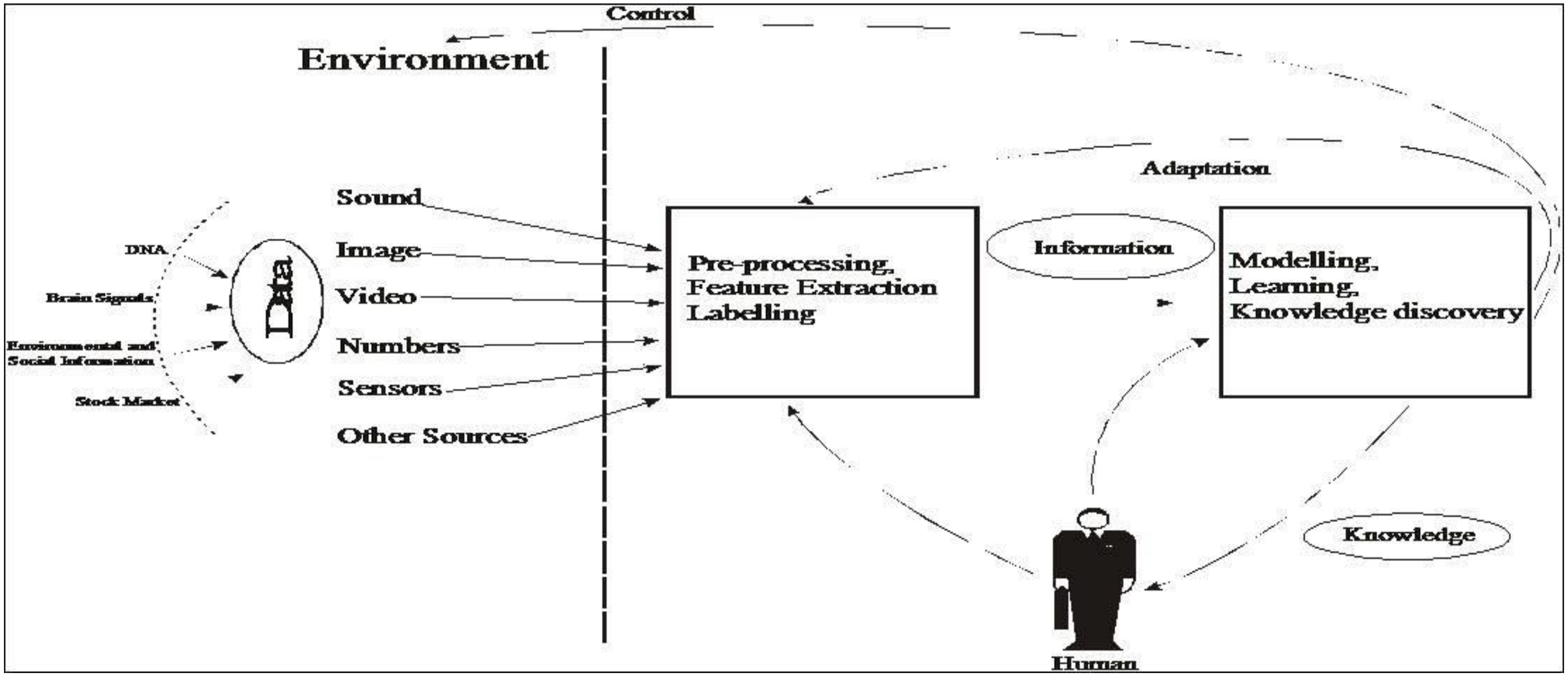
N.Kasabov, Evolving connectionist systems: The knowledge engineering approach, Springer, 2007 (first edition 2003), (AUT library, hard copy)

N.Kasabov, Foundations of neural network, fuzzy systems and knowledge engineering, MIT Press, 1996 (AUT library, hard copy)

1. Data Mining and Knowledge Engineering

- Data mining: Finding informative patterns and structured information from data
- Knowledge Engineering: Representing and Elucidating knowledge in intelligent information systems
- Intelligent systems – information systems that have features of intelligence, such as: learning, generalisation, pattern recognition, decision making, adaptation. Some of them acquire continuously such features over time , e.g. evolving connectionist systems.
- Learning systems: information systems that learn from data and improve their performance over time; e.g. neural networks or connectionist systems.
- Knowledge discovery – learning systems facilitate the extraction of new associations, rules, and relationships from data that are interpreted by humans. Rules can be propositional, fuzzy, temporal, etc.

Data collection, pre-processing, modeling, and knowledge discovery

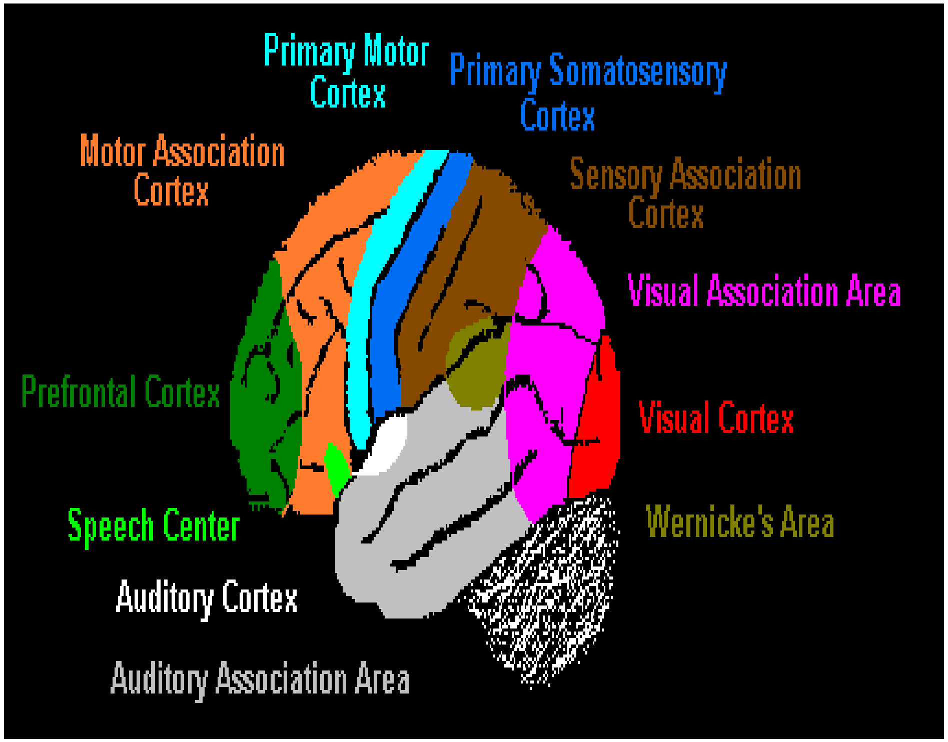


- Modelling complex processes is a difficult task: adaptation is needed
- Knowledge discovery
- A broad range of real-world applications

Model creation and model validation

- Training a model on training data
- Testing the model on test data
- Cross-validation (multiple model creation and testing)
- Leave-one –out
- "Un-biased" feature selection and modeling

2. Neural Networks: Inspiration from the brain and main principles



The brain:

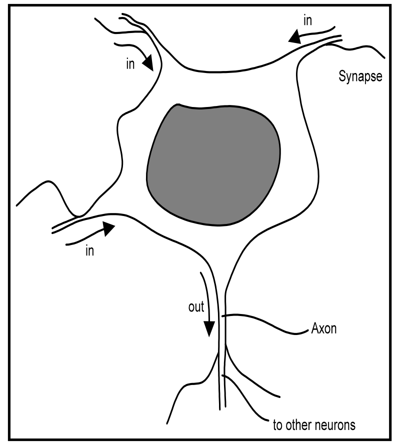
- The brain evolves through genetic "pre-wiring" and life-long learning at its different "levels"
- Evolving structures and functions
- Evolving features
- Evolving knowledge
- Local (e.g. cluster-based) learning and global optimisation
- Memory (prototype)-based learning, "traceable"
- Multimodal , incremental learning

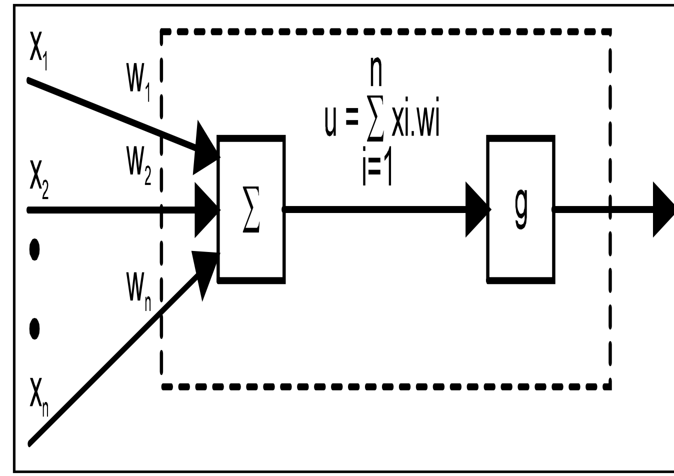
The challenge:

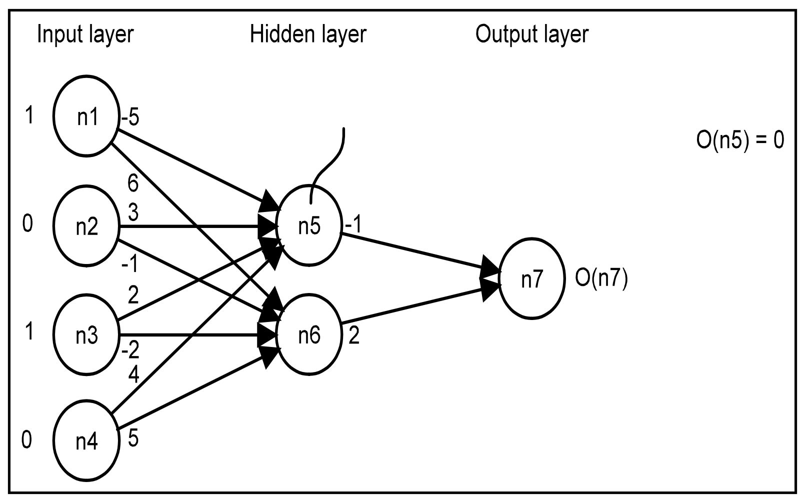
How do we achieve this in ANN and AI systems?

Artificial neural networks (ANN) (connectionist systems)

- ANN are computational models that mimic the nervous system in its main function of adaptive learning.
- ANN can *learn* from data and make *generalisation*
- ANN are *universal computational models*
- Software and hardware realisation of ANN
- The area of neurocomputing







NN for unsupervised learning

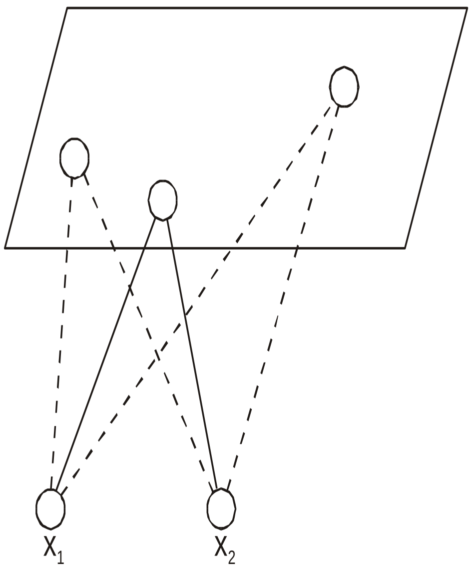
Unsupervised learning tasks:

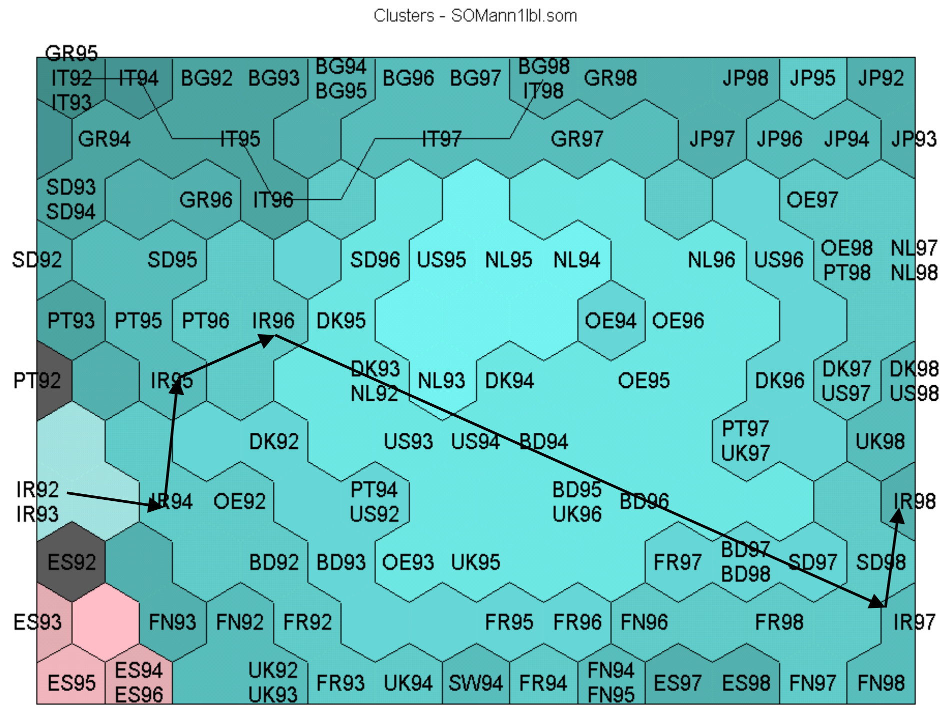
- Clustering
  - Discovering groupings (clusters) of data represented by:
    - Cluster centres
    - Membership degree of each sample to each cluster
    - Exact and fuzzy clustering
- Vector Quantisation
  - Mapping data vectors from into a smaller dimensional space
- Prototype Learning
  - Similar to clustering, but instead of cluster centers, prototypes of data points are found to represent the data to some degree of accuracy

Self Organising Maps (SOMs)

- Teuvo Kohonen, TU Helsinki
- Belong to vector quantisation methods
- Each output neuron specializes during the training to react to similar input vectors from a group (cluster) of the input space
- Neurons in output layer are competitive
- SOMs preserve similarity between input vectors from the input space as *topological closeness of neurons* in the output space represented as a topological map.

SOMs





A schematic diagram of a simple, hypothetical two-input, 2D output SOM system

NN for supervised learning. MLP

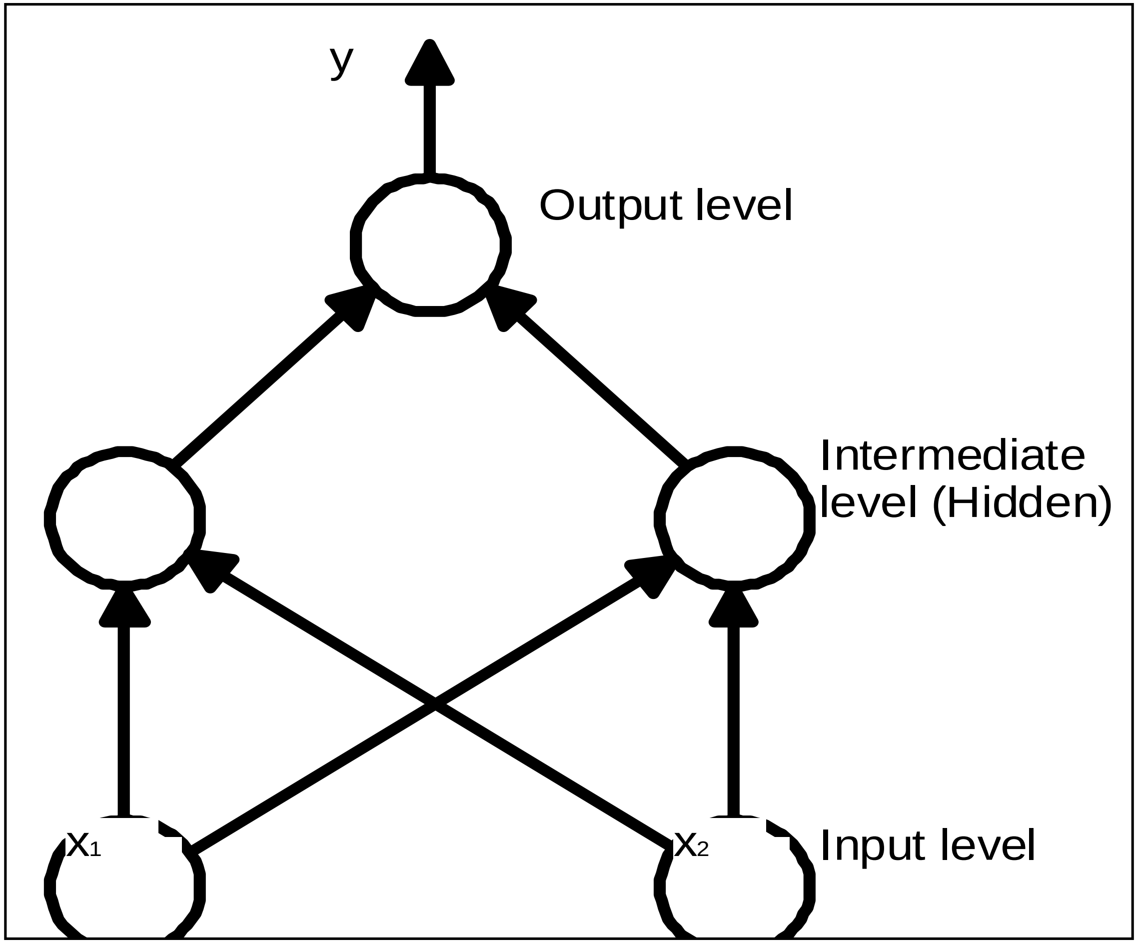
- The learning principle is to provide the input values and the desired output values for each of the training examples.

- The neural network changes its connection weights during training.

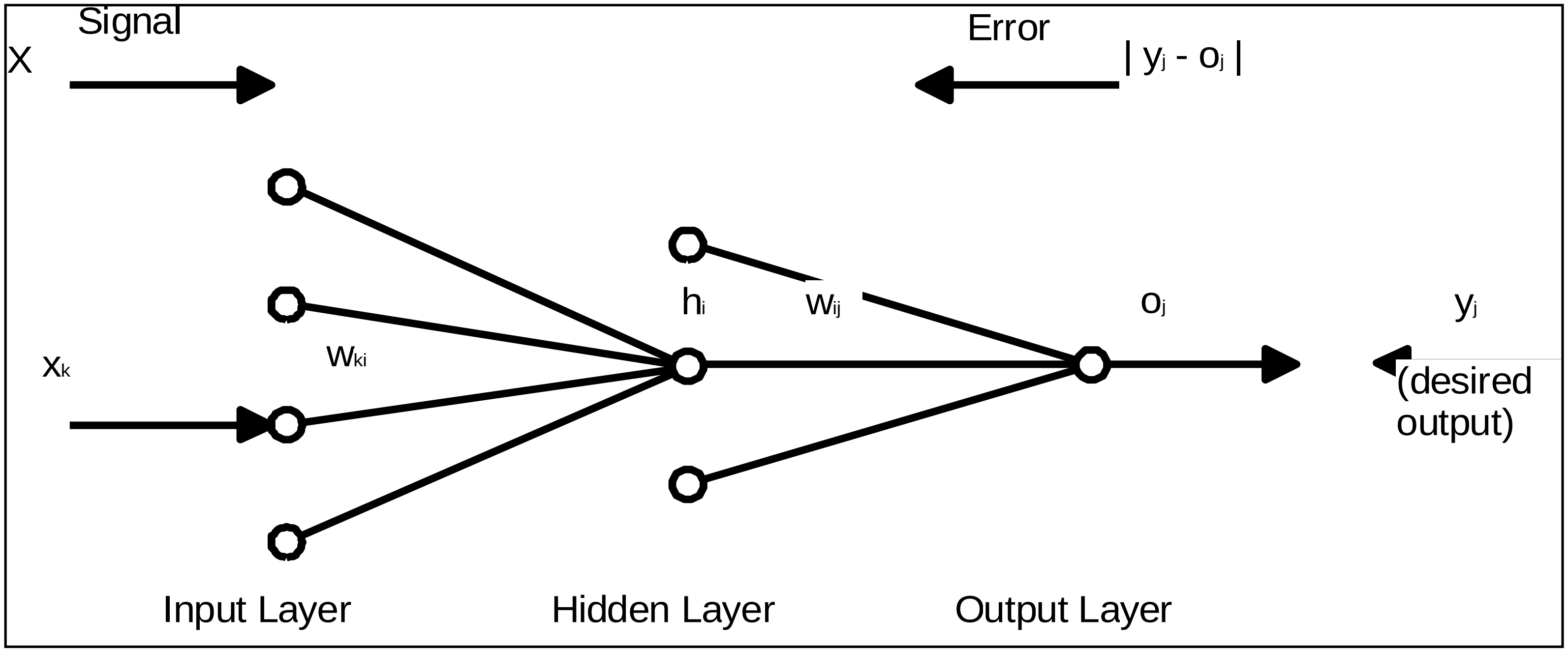
- Calculate the error:
  - training error - how well a NN has learned the data
  - test error - how well a trained NN generalises over new input data.

MLP and the backpropagation algorithm (Rumelhart et al, 1986)

- Solving the problem of linear non-separability



MLP and the backpropagation algorithm



MLP and the backpropagation algorithm

Forward pass:

BF1. Apply an input vector **x** and its corresponding output vector **y** (the desired output).

BF2. Propagate forward the input signals through all the neurons in all the layers and calculate the output signals.

BF3. Calculate the Err_j for every output neuron j as for example:
$Err_j = y_j - o_j$, where y_j is the jth element of the desired output vector **y**.

Backward pass:

BB1. Adjust the weights between the intermediate neurons i and output neurons j according to the calculated error:
$\Delta w_{ij}(t+1) = lrate.\ o_j(1 - o_j).\ Err_j.\ o_i + momentum.\ \Delta w_{ij}\ (t)$

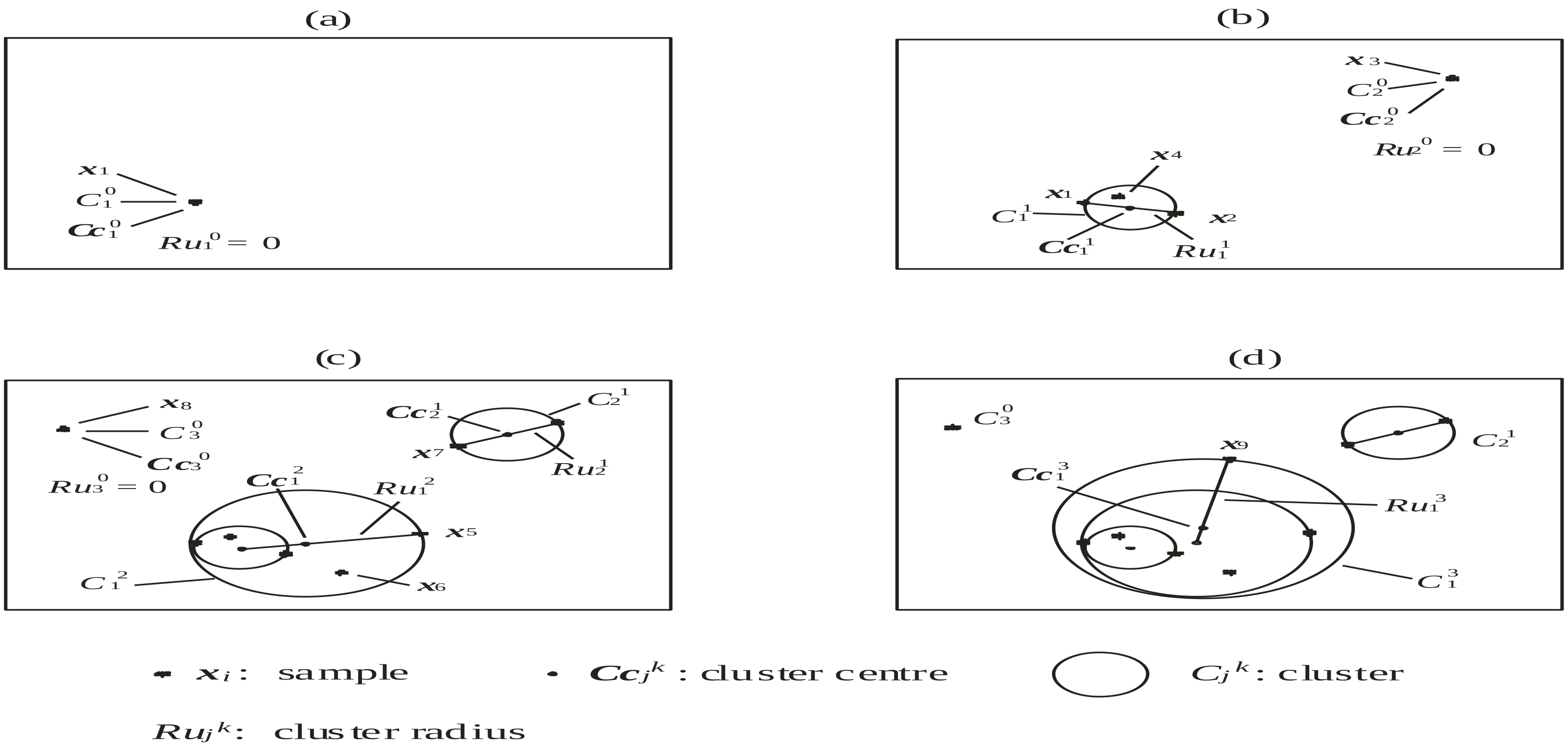
BB2. Calculate the error Err_i for neurons i in the intermediate layer:
$Err_i = \sum Err_j.\ w_{ij}$

BB3. Propagate the error back to the neurons k of lower level:
$\Delta w_{ki}(t+1) = lrate.o_i(1 - o_i).\ Err_i.x_k + momentum.\ \Delta w_{ki}(t)$

3. Evolving Connectionist and Hybrid Neural Networks

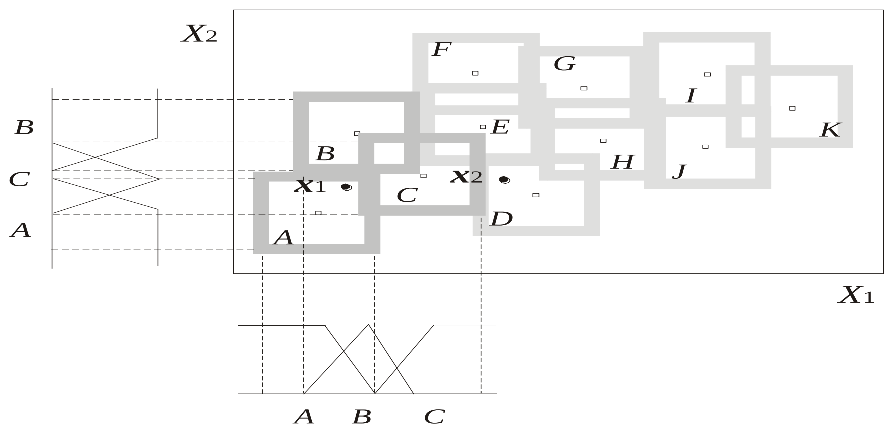
Evolving clustering

- Input stream data
- Incremental clustering: Every new sample is assigned to the closest existing cluster or a new cluster is created based on distance measure.
- Local learning based on evolving clustering

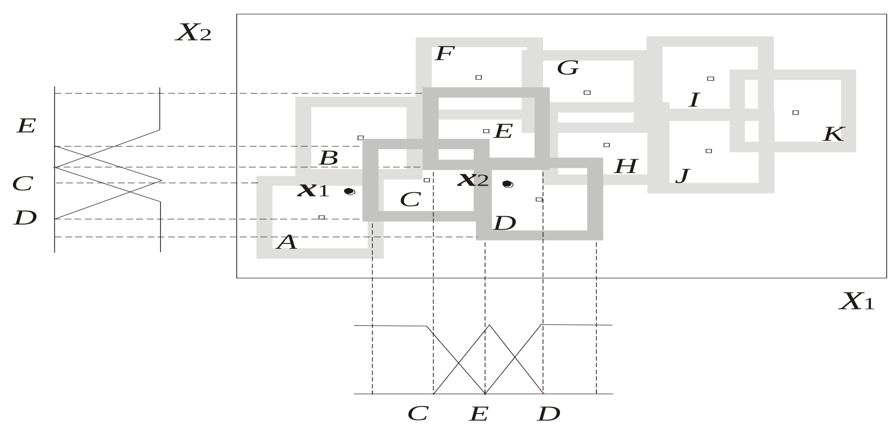


Dynamic Evolving Neuro-Fuzzy Inference System (DENFIS) for Supervised Learning (DENFIS, Kasabov and Song, 2002, IEEE Tr Fuzzy Systems, 800citations)

(a) Fuzzy rule group 1 for a DENFIS



(b) Fuzzy rule group 2 for a DENFIS



DENFIS algorithm:

(1) Learning:

- Unsupervised, incremental clustering.
- For each cluster there is a Takagi-Sugeno fuzzy rule created: IF x is in cluster Cj THEN yj = fj (x),

where: $y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_q$

- Incremental learning of the function coefficients and weights of the functions through least square error

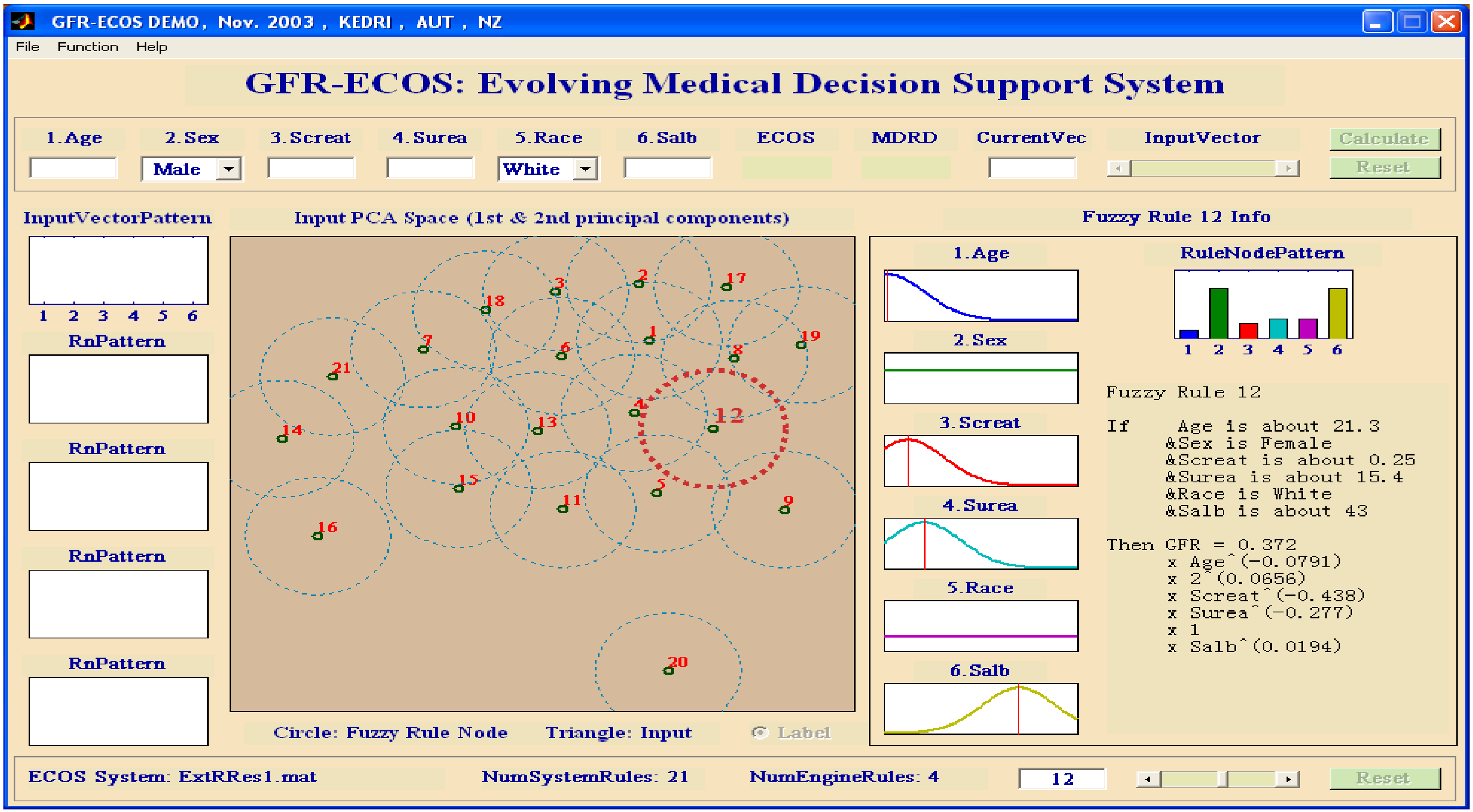
(2) Fuzzy inference over fuzzy rules:

- For a new input vector x = [x1,x2, … , xq] DENFIS chooses m fuzzy rules from the whole fuzzy rule set for forming a current inference system.
- The inference result is:

$$y = \frac{\sum_{i=1,m} [\, \omega_i f_i (x_1, x_2, \ldots, x_q)]}{\sum_{i=1,m} \omega_i}$$

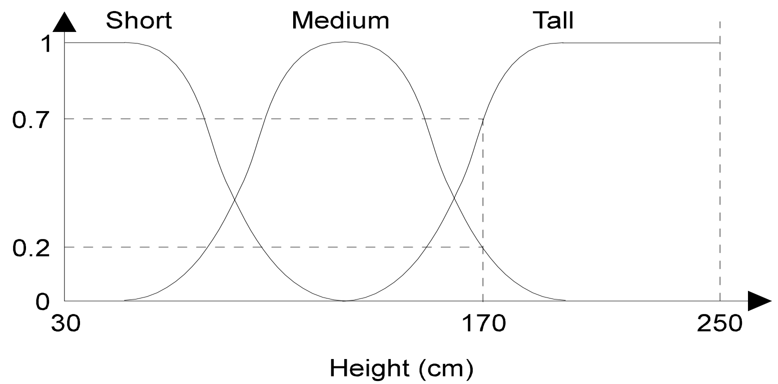
Application example of DENFIS: Renal Function Evaluation System

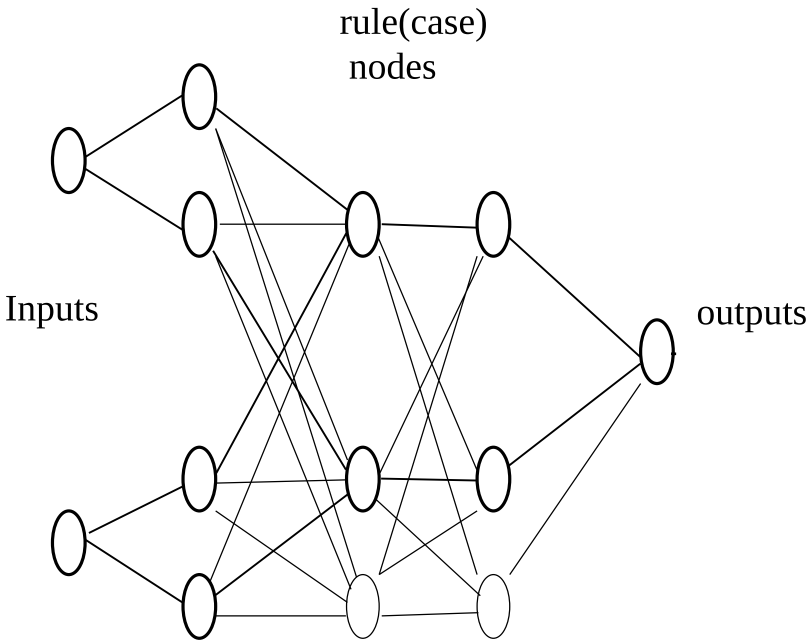
Marshal, Song, Ma, McDonell and Kasabov, Kidney International, May 2005)



Evolving Fuzzy Neural Network (EFuNN) for supervised learning

- As a general case, input and/or output variables can be non-fuzzy (crisp) or fuzzy
- Fuzzy variables
- Example of three Gaussian MF





- EFuNN, N. Kasabov, IEEE Tr SMC, 2001

- Incremental, supervised clustering

- Weights change based on *Euclidean distance* between input vectors and prototype nodes:

  $\Delta \mathbf{w} = lrate * E(\mathbf{x}, Rn)$

- Evolving Zadeh-Mamdani or Takagi-Sugeno fuzzy rules as knowledge representation

Knowledge manipulation in EFuNN

- Important for an ECOS not only to learn in lifelong learning mode, but also to “explain” at any time the essence/knowledge that the system has acquired
- Rule Insertion and Extraction
  - Fuzzy or exact rules can be inserted and extracted at any phase of the learning process

Example:

Rule 1 :

IF input [1] is (Small 0.46) and (Medium 0.540) and input [2] is (Large 0.809) THEN output is (Large 0.685); Radius of the receptive field 0.106; Number of accommodated examples = 2

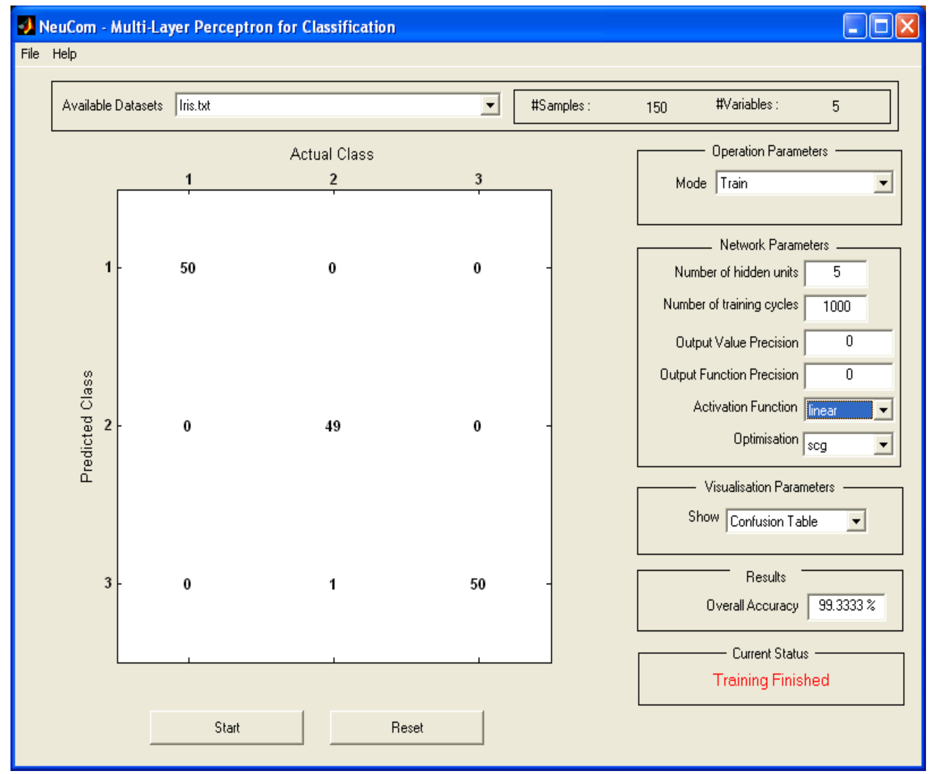
Rule 2:

IF input [1] is (Medium 0.527) and (Large 0.473) and input [2] is (Small 0.461) and (Medium 0.539)

THEN output is (Small 0.496) and (Medium 0.504); Radius of the receptive field = 0.124; Number of examples = 5

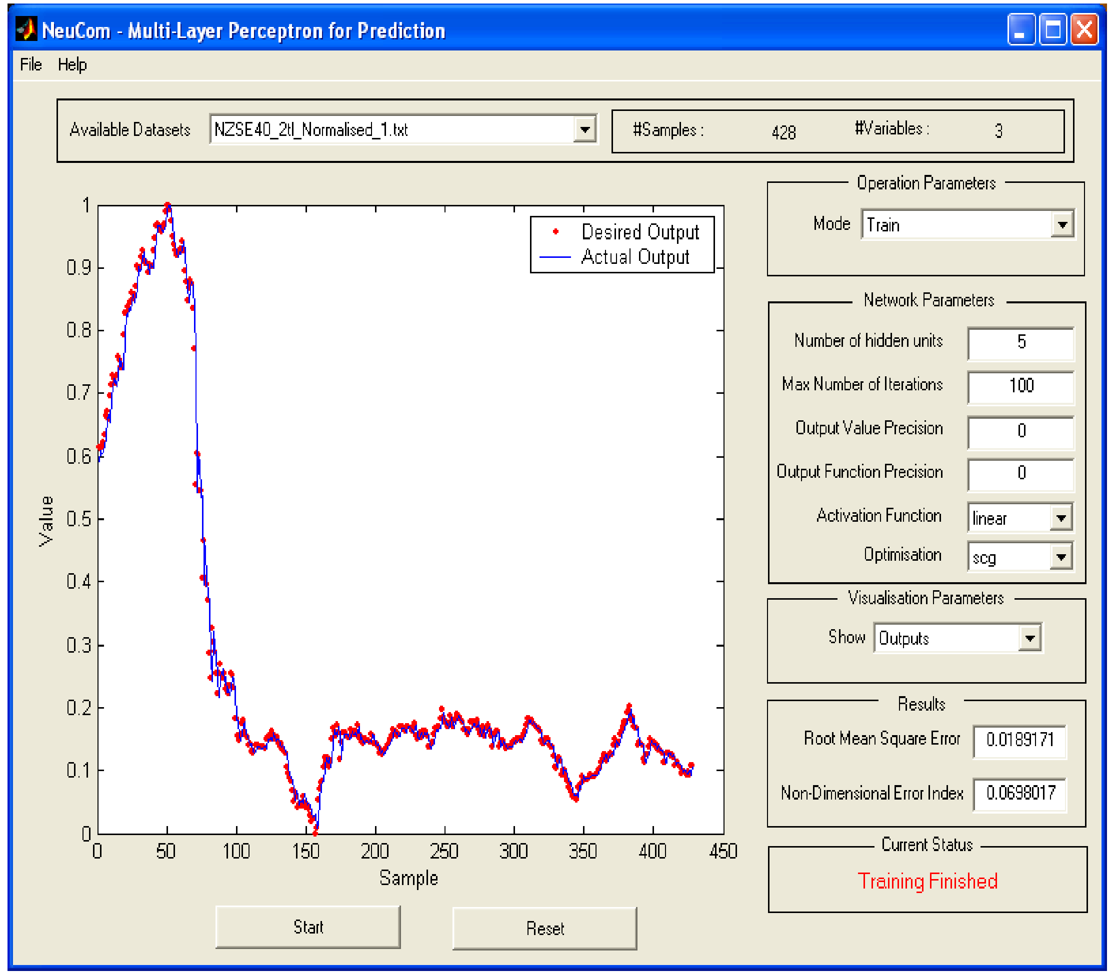
4. Applications of NN and ECOS
Classification

- The outputs are class lables
- Calculating the confusion matrix:
  - True-positive (sensitivity)
  - True negative (specificity)
- Iris data
- Comparison between different NN methods in NeuCom



Prediction

- Time series prediction
- Choosing the time-lags and the features
- Case studies using NeuCom
- Training on data
- Model verification
- Gas furnice time series prediction
- Stock index time series prediction
- Comparison between different NN methods in NeuCom



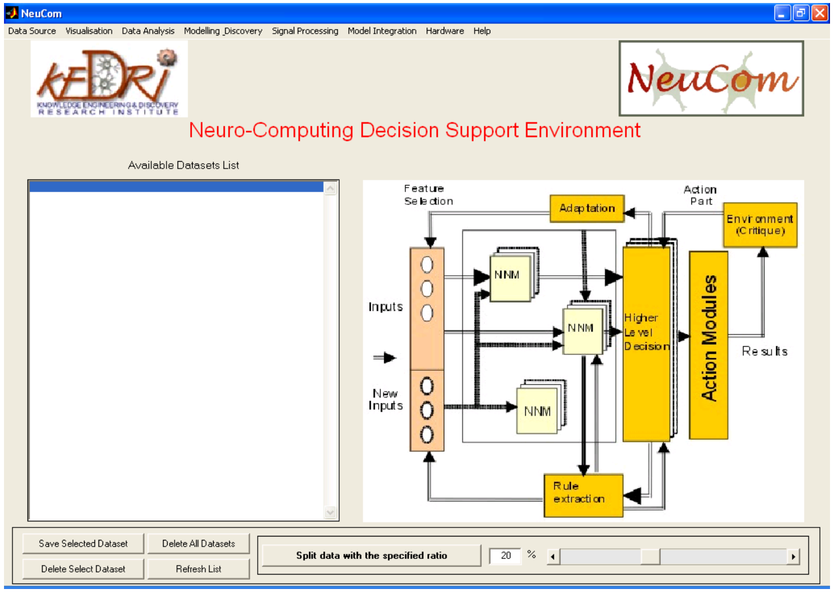
Other applications

- Bioinformatics and biomedical applications (chapter 8 from ECOS and SHBNI)
- Neuroinformatics, e.g. applications for brain study, BCI, etc. (chapter 9 from ECOS and SHBNI)
- Signal processing: speech, image (chapters 10-13 rom ECOS)
- “Artificial nose”, “artificial tongue” (SHBNI)
- Decision support systems, Ecology, Environment, Bio-protection, robotics (chapter 14 from ECOS and SHBNI)

5. Introduction and demonstration of NeuCom:

A Software Environment for NeuroComputing, Data Mining and Intelligent System Design (www.theneucom.com)

- A generic environment, that incorporates 60 traditional and new techniques for intelligent data analysis and the creation of intelligent systems
- Methods for feature selection
- Methods for classification
- Methods for prediction
- Methods for knowledge extraction
- Fast data analysis and visualisation
- Fast model prototyping
- A free copy available for education and research from: www.theneucom.com
- Adopted in 20 universities all over the world and in research laboratories



6. References

- N.Kasabov, Evolving connectionist systems, Springer Verlag, 2007
- SHBNI - N.Kasabov (ed) Springer Handbook of Bio-/Neuroinformatics, 2014 (available free as a .pdf from the AUT Library as an eBook and also as a hard copy)