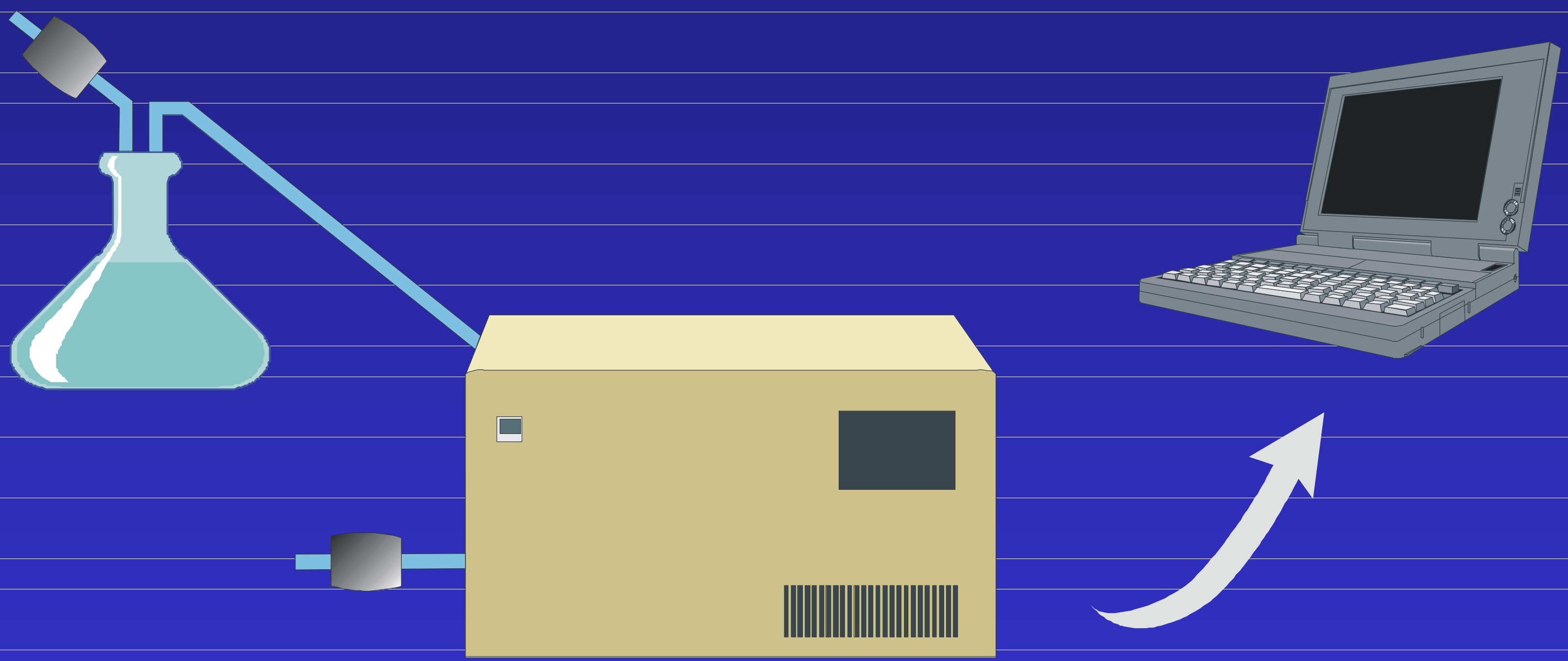
Food Quality Assurance Applying a Sophisticated Neural Network to Olfactory Signals

Roland Linder, Siegfried J. Pöppl

- Methods
- Data
- Results
- ACMD – what does it mean?
- Conclusions

Portable Electronic Nose (PEN)



Sample

Signal Detection Unit

Signal Processing Unit
(Classification)

We are not the first ones who use an EN

- Marsili, RT: Shelf-life prediction of processed **milk** by solid-phase microextraction, mass spectrometry, and multivariate analysis. J Agric Food Chem 2000;48(8):3470-5.
- Blixt Y, Borch E: Using an electronic nose for determining the spoilage of vacuum-packaged **beef**. Int J Food Microbiol 1999;46(2):123-34.
- Young H, Rossiter K, Wang M, Miller M: Characterization of Royal Gala **apple** aroma using electronic nose technology-potential maturity indicator. J Agric Food Chem 1999;47(12):5173-7.

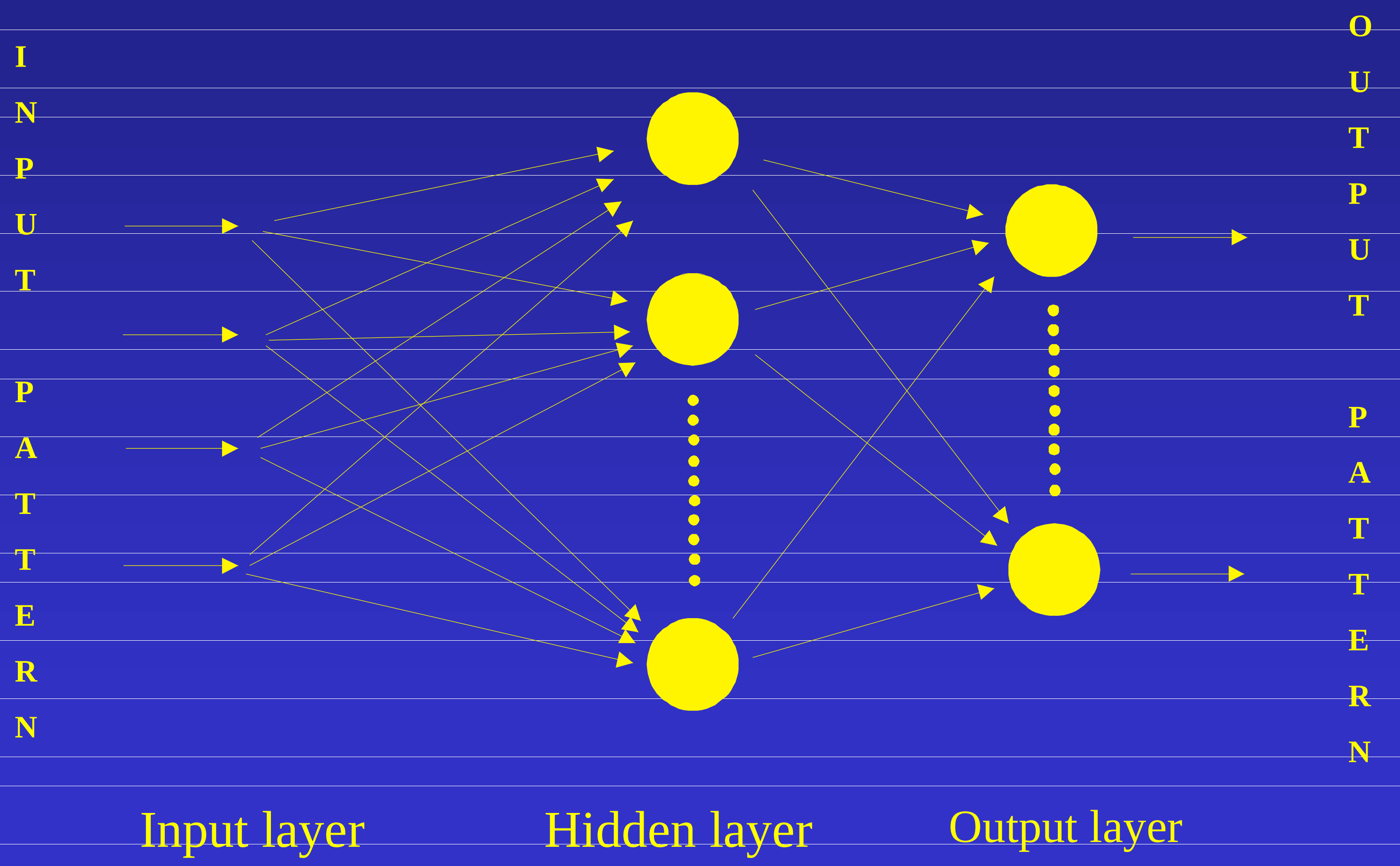
We are not the first ones who use an Artificial Neural Network (ANN)

- Boilet P, Hines EL, John S, Mitchell J, Lopez F, Gardner JW, Llobet E, Hero M, Fink C, Gongora M: Detection of eye **bacteria** causing eye infections using a **neural network** based electronic system. In Gardner JW, Persaud KC (eds.): **Electronic Noses** and Olfaction, IOP Publishing Ltd, Bristol, 2000, 189-196.
- Llobet E, Hines EL, Gardner JW, Franco S: Non-destructive **banana** ripeness determination using a **neural network**-based **electronic nose**. Meas Sci Technol 1999;10(6):538-548.
- Hines EL, Llobet E, Gardner JW: **Neural network** based **electronic nose** for **apple** ripeness determination. IEE Electronic Letters 1999;35(10):821-823.

What is the benefit of using an ANN ?

- Non-linearities
- Dependencies between variables
- No assumptions of any well-defined distributions

How does an ANN work?



Learning algorithms we used

- Standard Backpropagation (**BPN**)

  [Rumelhart, Hinton & Williams 1986]
- Resilient Propagation (**RPROP**)

  [Riedmiller & Braun 1993]
- Adaptive Propagation (**APROP**)

  [Linder, Wirtz & Pöppl 2000]

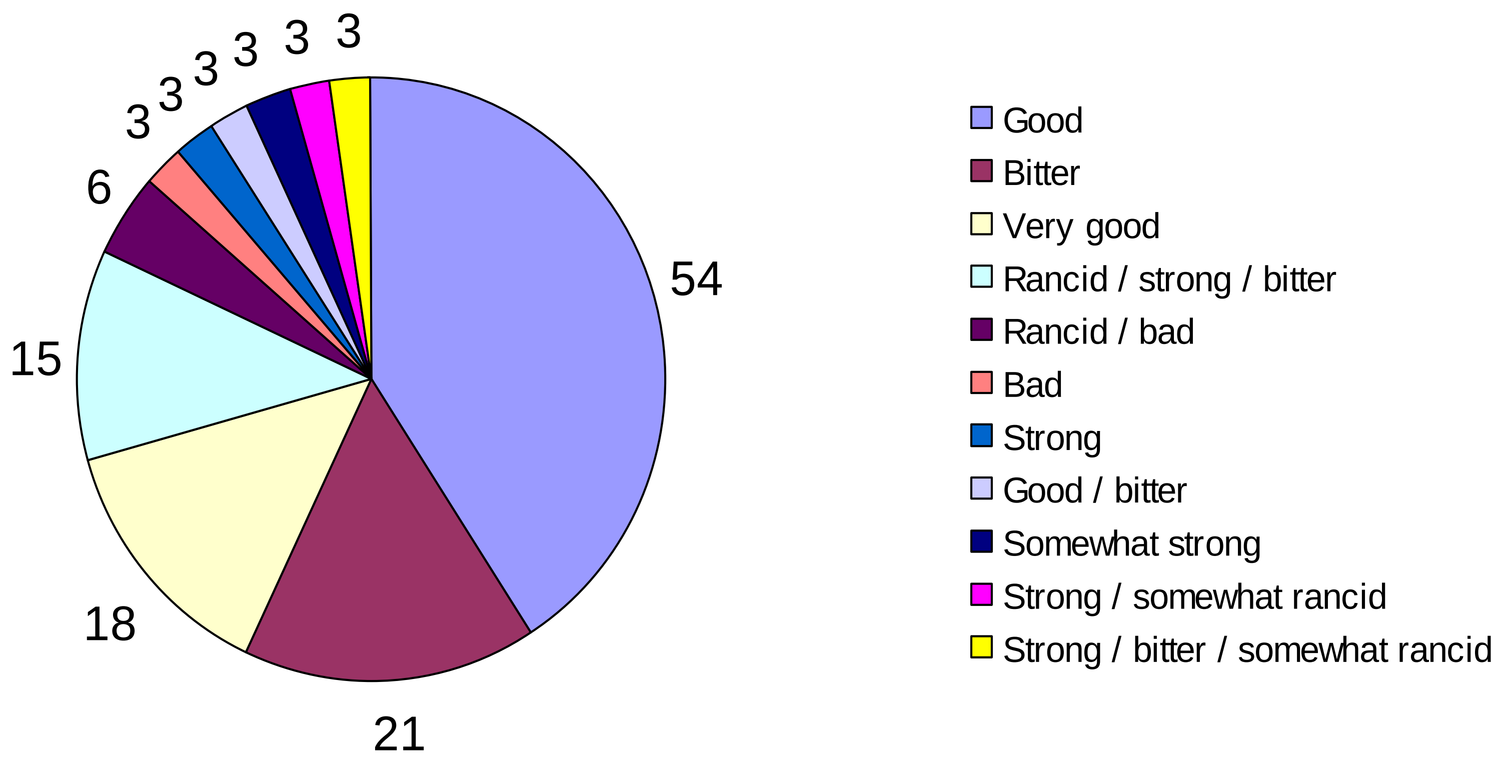
Standard Backpropagation (BPN)

$$\Delta w_{jk}^{new} = \eta \cdot \frac{\partial E}{\partial w_{jk}} + \alpha \cdot \Delta w_{jk}^{old}$$

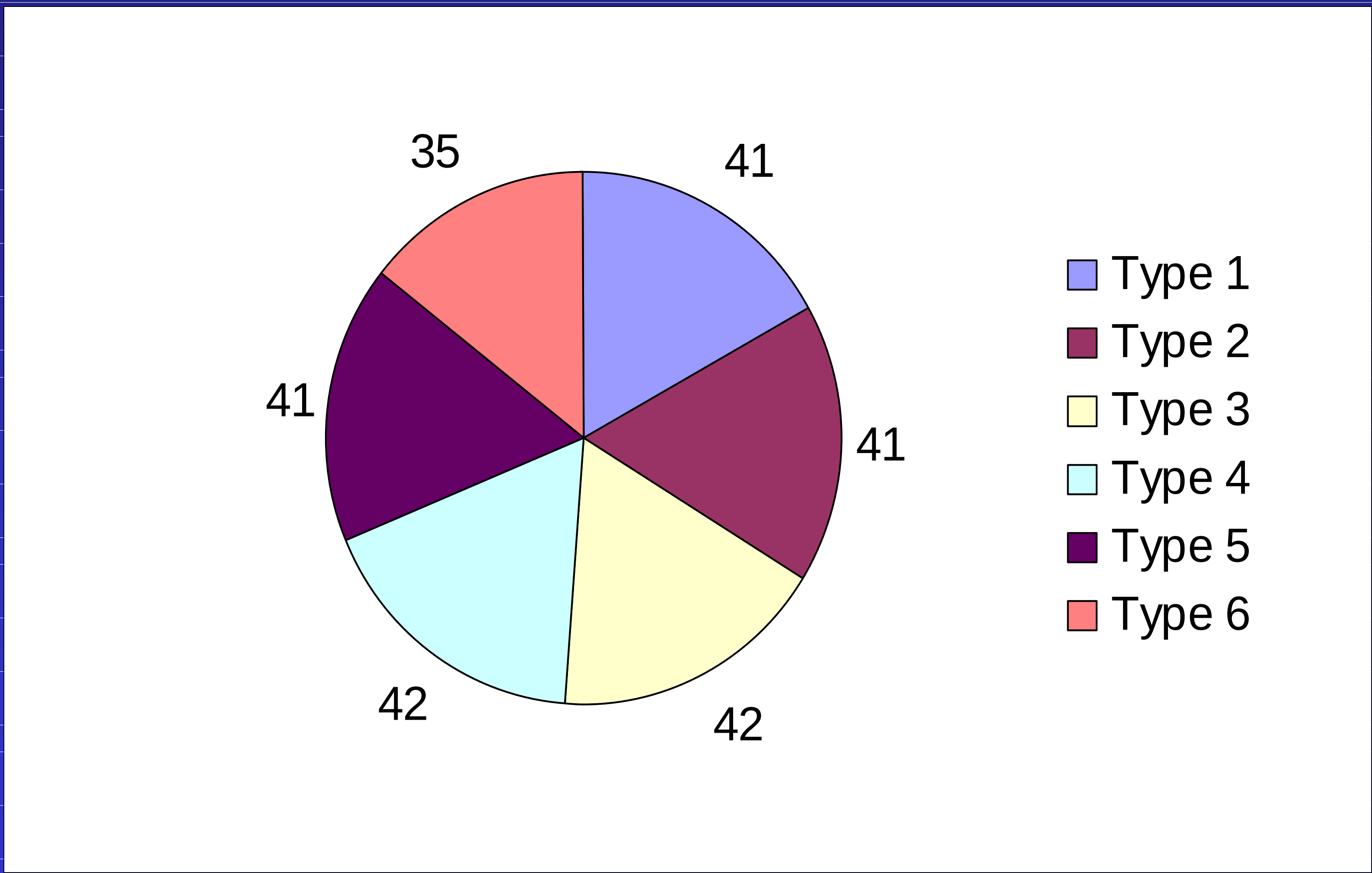
Resilient Propagation (RPROP)

$$\Delta w_{ij} = \begin{cases} -\Delta_{ij} & \text{, if } \dfrac{\partial E}{\partial w_{ij}} > 0 \\ +\Delta_{ij} & \text{, if } \dfrac{\partial E}{\partial w_{ij}} < 0 \\ 0 & \text{, otherwise} \end{cases}$$

Classifying Edible Oil

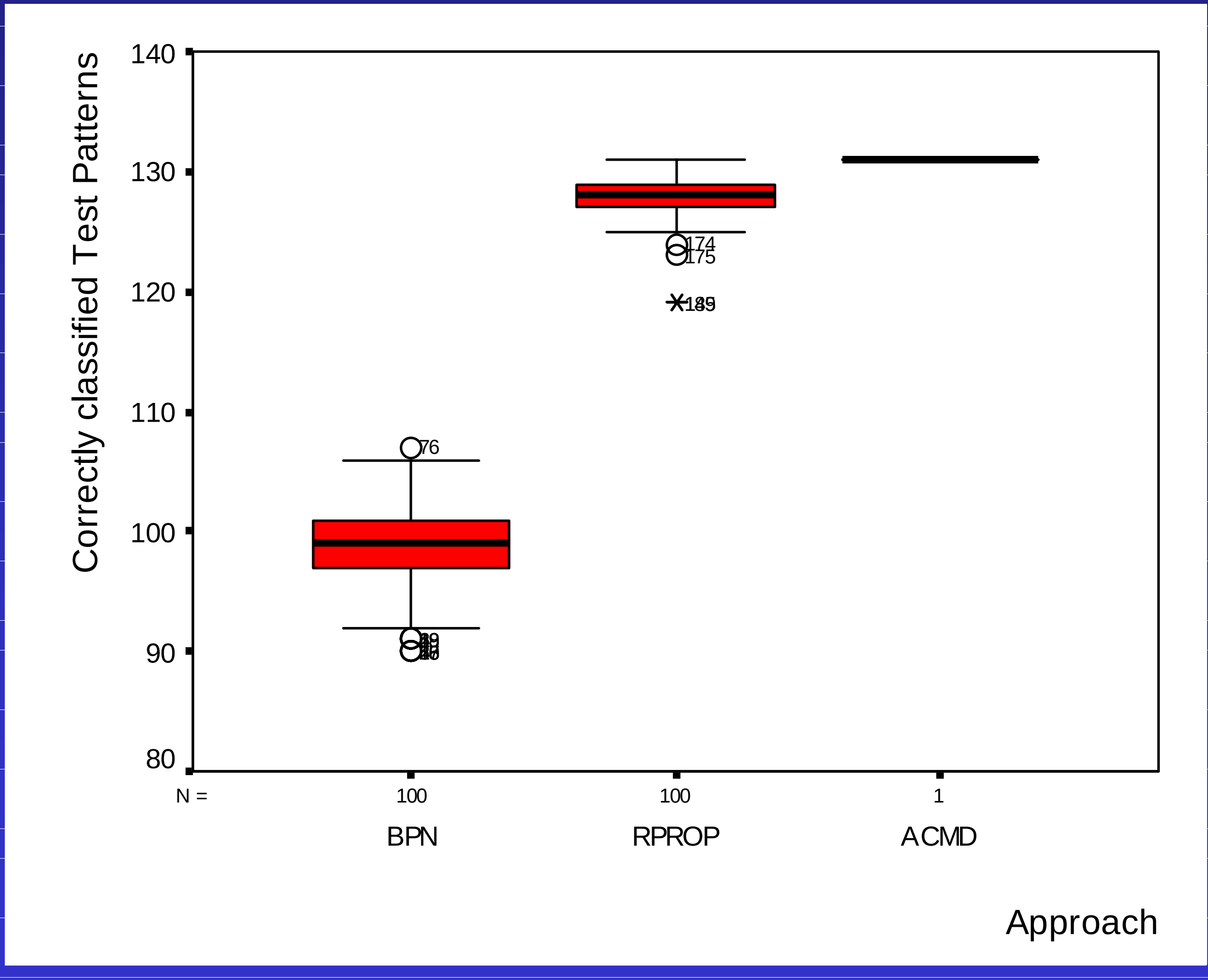


Classifying Honey

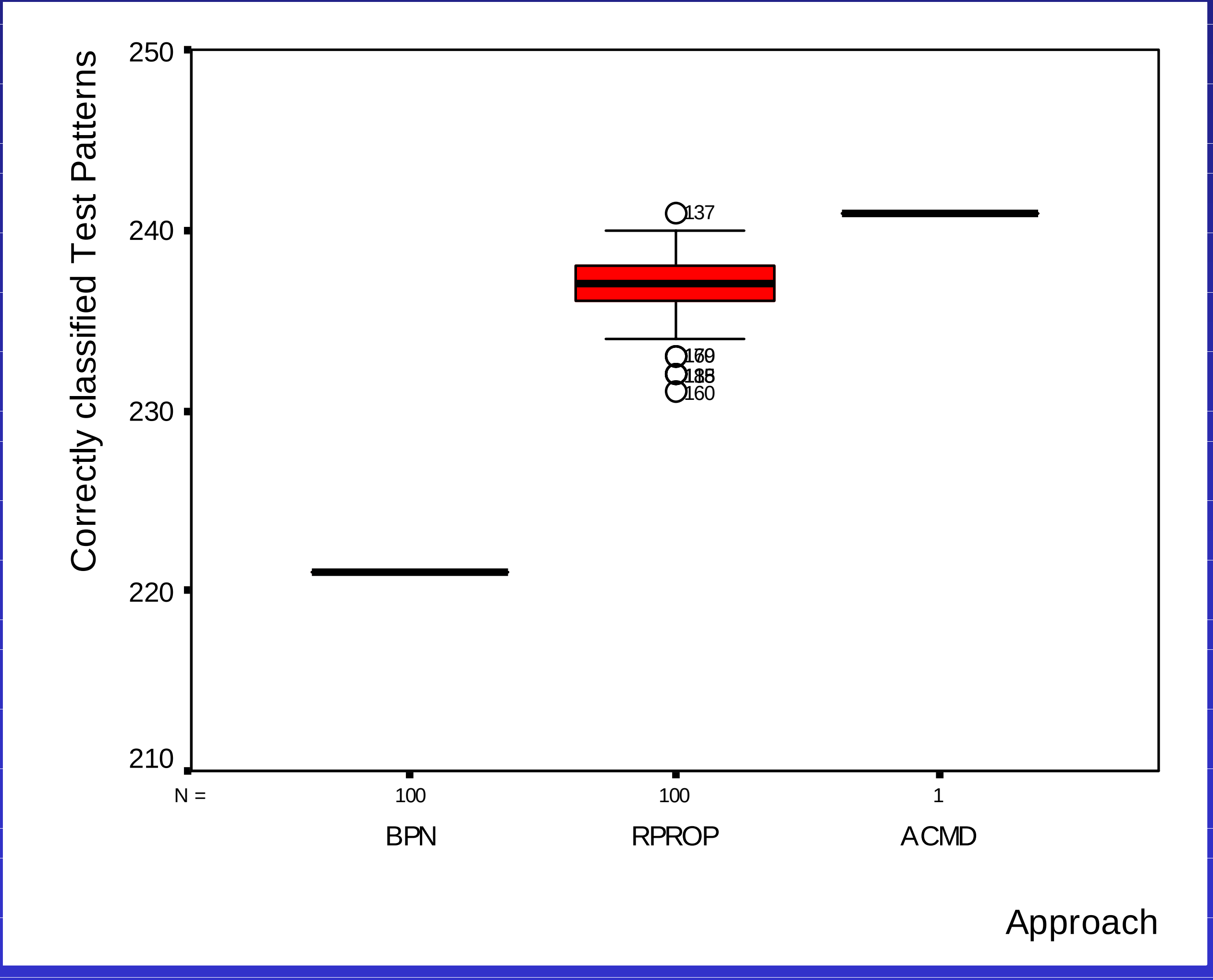


The Neural Network

Results: Edible Oil



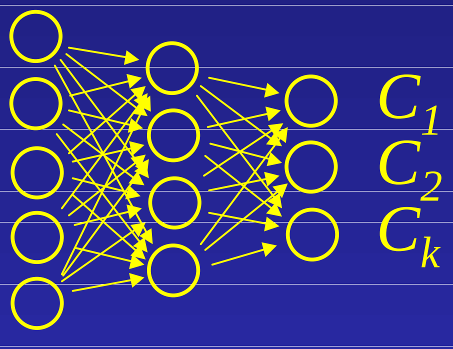
Results: Honey



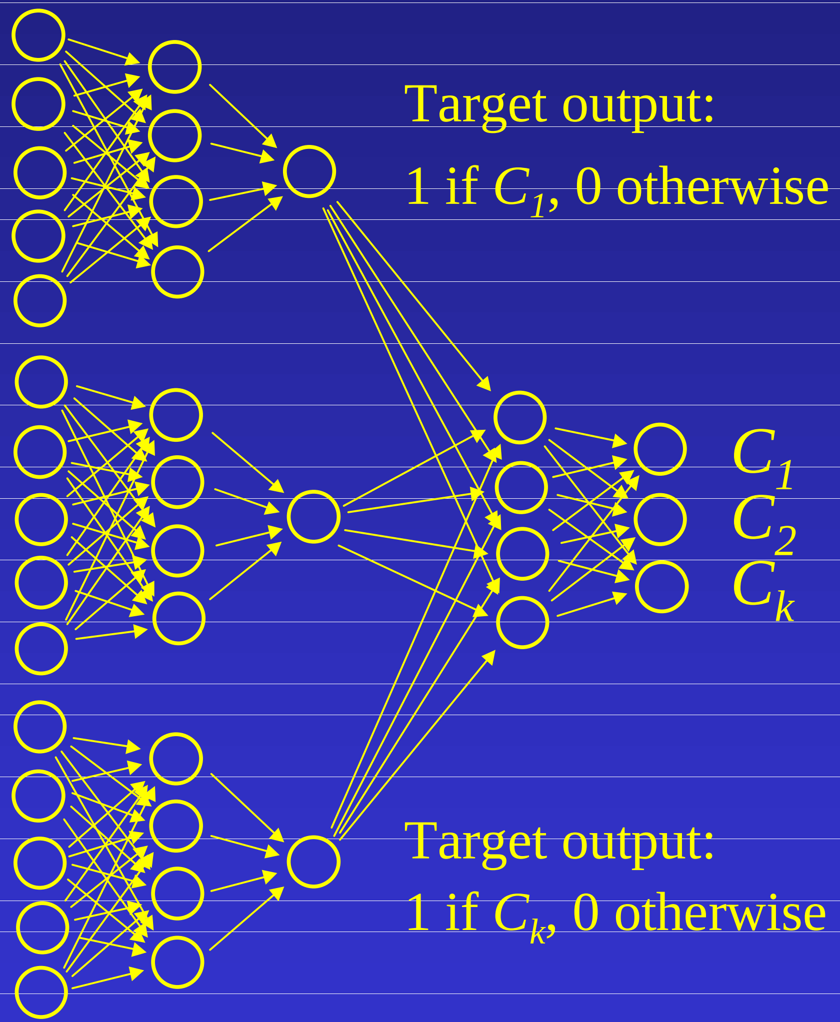
Strategies – an Overview

Strategy	Generalisation	Speed	No tuning
Multi-neural-network	x	(x)	
Refining the Target output	x		
Early stopping	x		x
Avoiding weight sets from oscillating phases	x		
Oversizing the network	x		x
Network ensemble	x		
Exclusion Strategy	x		
Adaptive propagation		x	
Modifying the error function	x		
Modifying the derivation	x		
Varying the learning rate	x		x
Stop oscillating and unsuccessful learning		x	
Optimising the software		x	
Approximation of the exponential function		x	

Multi-Neural-Network



k-class problem



k two-class problems

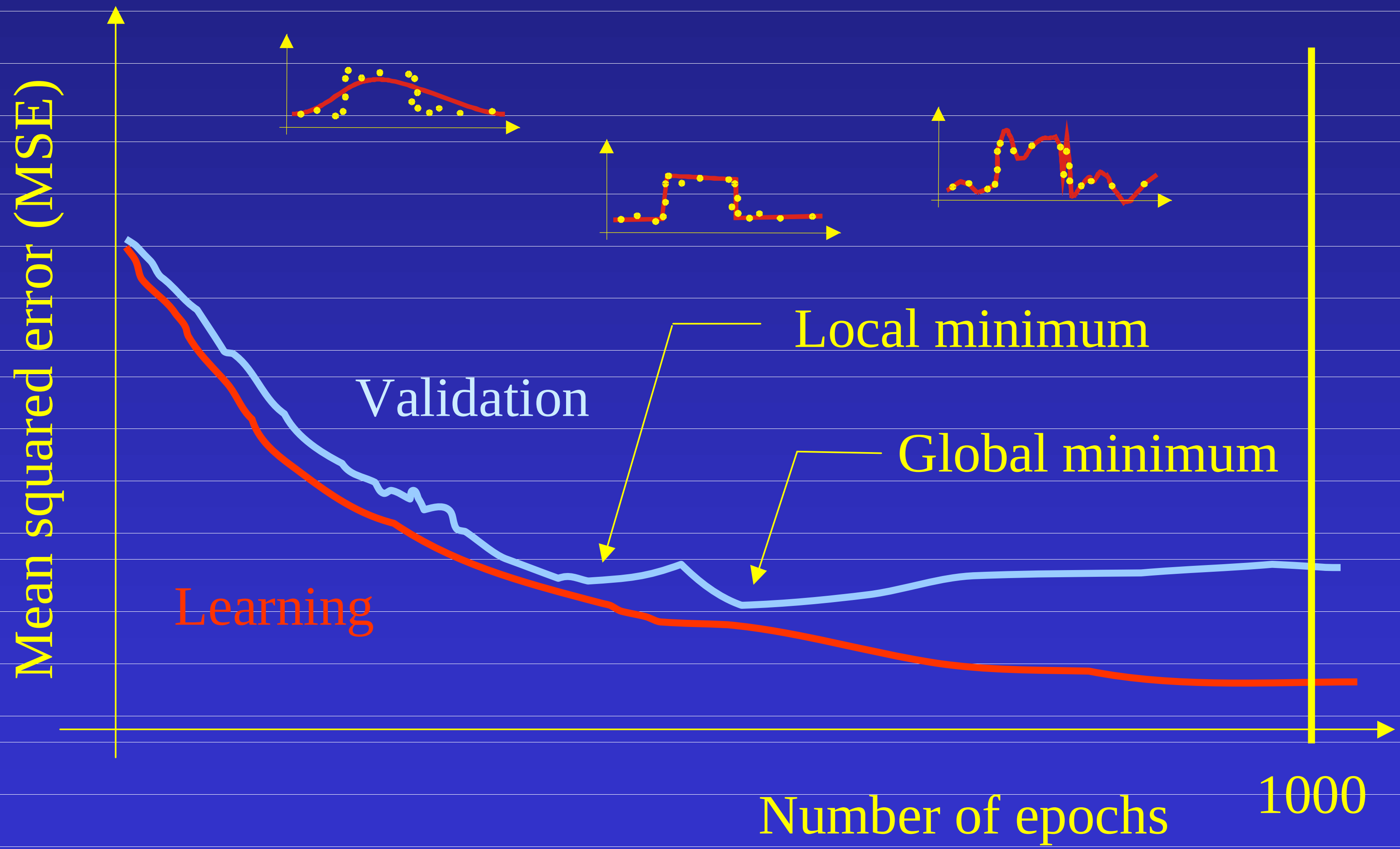
Multi-Neural-Network

- Simple functions can be better learned than complex functions.
- Learning may be faster and parallel computing is possible.
- Weight analysis is easier for modular networks.
- Better generalisation performance due to the cascaded architecture.

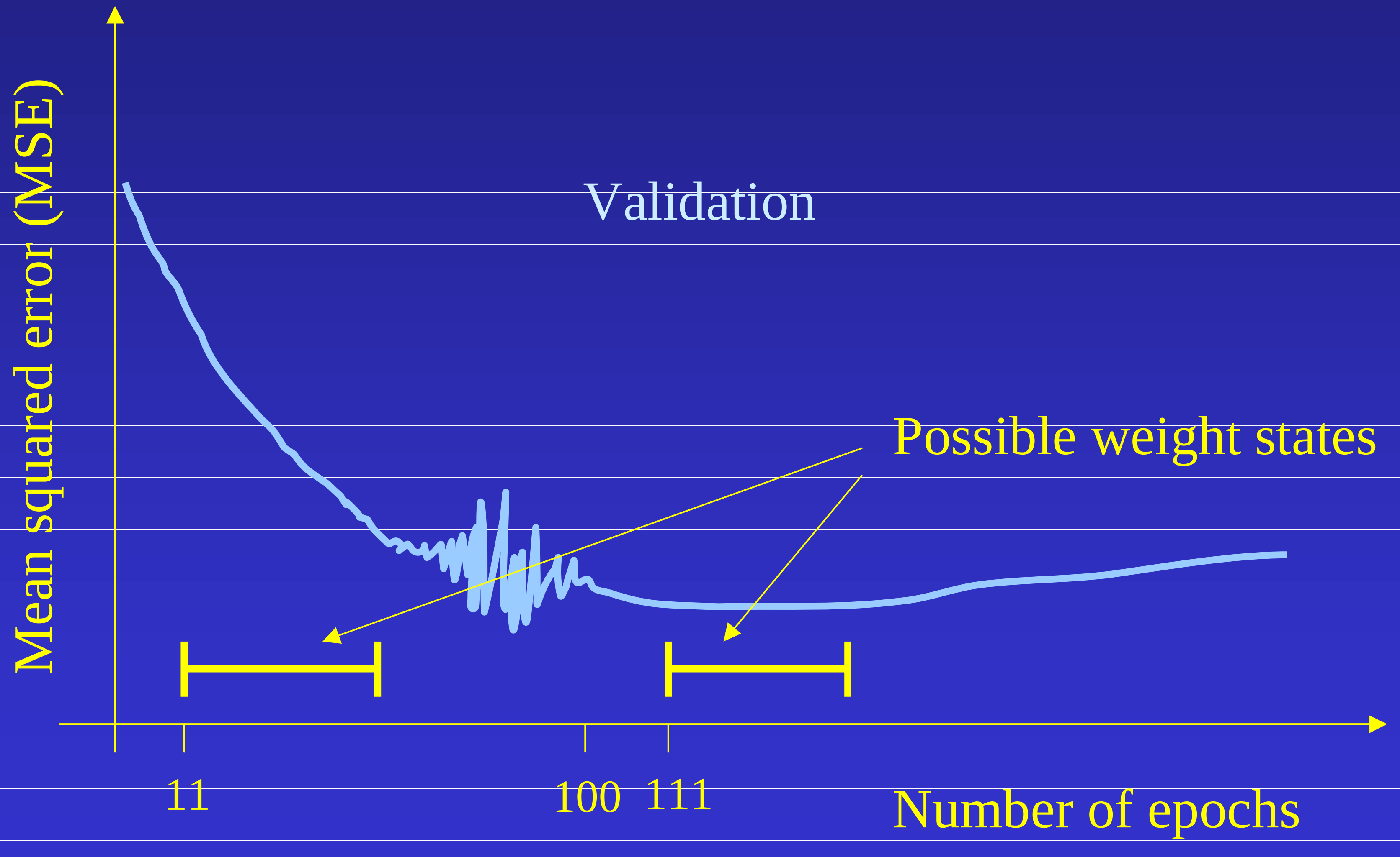
Refining the Target Output

Target output (1-out-of-C code)	Real output after 1000 epochs	Refined target output	| Target – real output |
1	0.7	1	0.3
0	0.3	0.3	0
0	0.1	0.1	0
0	0.2	0.2	0
0	0.6	0	0.6
0	0.1	0.1	0
0	0.1	0.1	0

Early Stopping



Avoiding Weight Sets from Oscillating Phases



Oversizing the Network

Lawrence, Giles, Tsoi: „We further support the observation that larger networks can produce better training and generalization error.“

Caruana: „Large networks rarely do worse than small networks.“

Hassoun, Cherkassky, Hanson, Oja, Sarle, Sudjianto: „It is clear that too many parameters in some nonparametric models can be grievous, however with many Neural Networks, more parameters can actually improve things.“

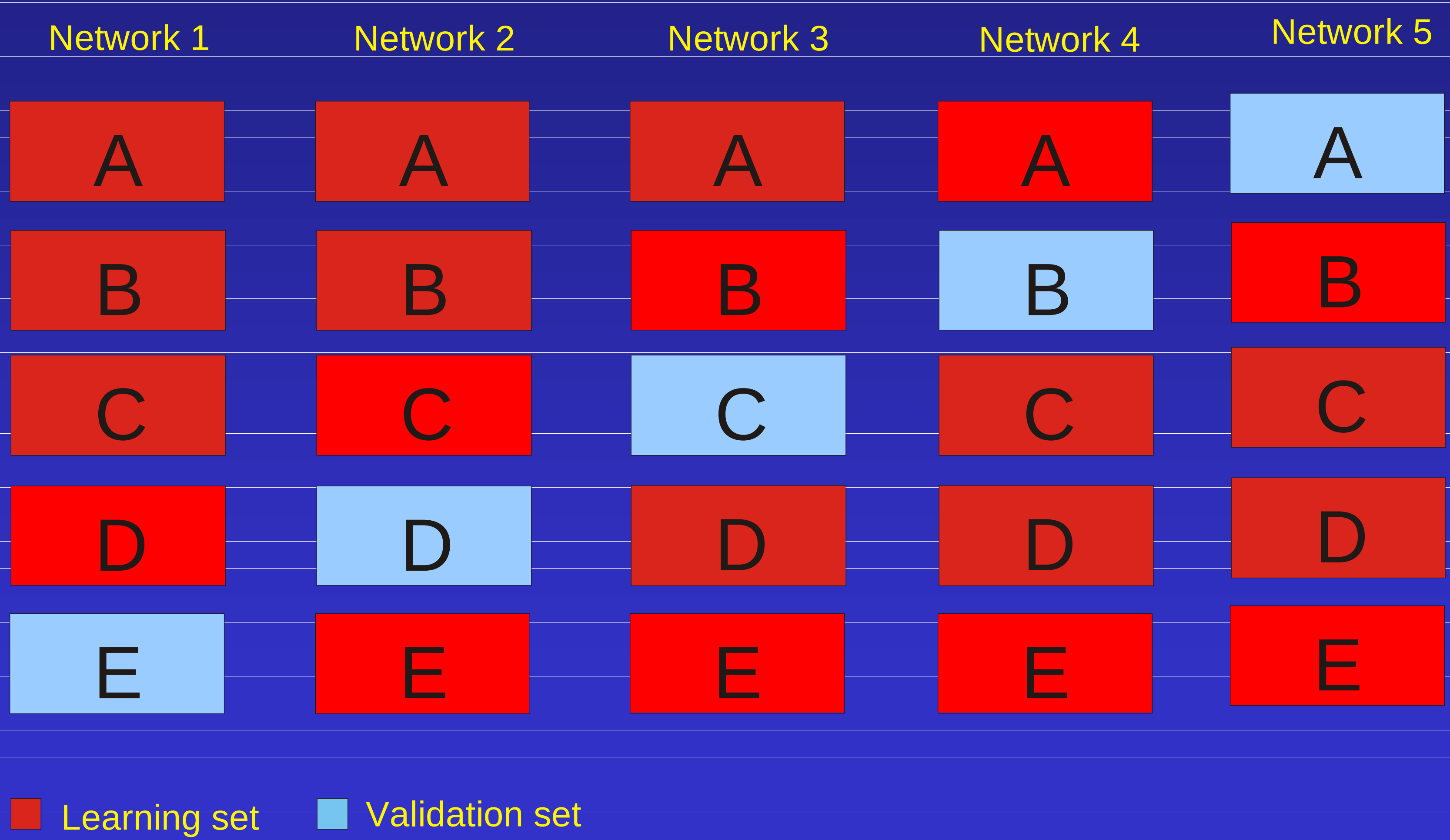
Kröse, van der Smagt: „The addition of extra parameters can decrease the chance of becoming stuck in local minima or on ‚plateaus‘, etc.“

Bartlett: „Neural networks often perform successfully with training sets that are considerably smaller than the number of network parameters.

Weigend: „Large networks generalize better than small ones.“

Rumelhart: „Adding a few more connections creates extra dimensions in weight-space and these dimensions provide paths around the barriers that create poor local minima in the lower dimensional subspaces.“

Network Ensemble



Exclusion Strategy

1 test pattern (with unknown class assignment)

Ensemble, trained with *n* classes

1) 0.9 2) 0.2 **3)** 0.8 4) ...

Ensemble, trained with class 1 + 2

Ensemble, trained with class 1 + 3

• • •

Ensemble, trained with class 1 + n

• • •

1) 0.86 **3)** 0.88

Result: Class 3

Adaptive Propagation (APROP)

$$S_{l,n} = \frac{1}{\sqrt{\sum_{i=1}^{p} \delta^2_{l,n,i}}}$$

Modifying the Error Function

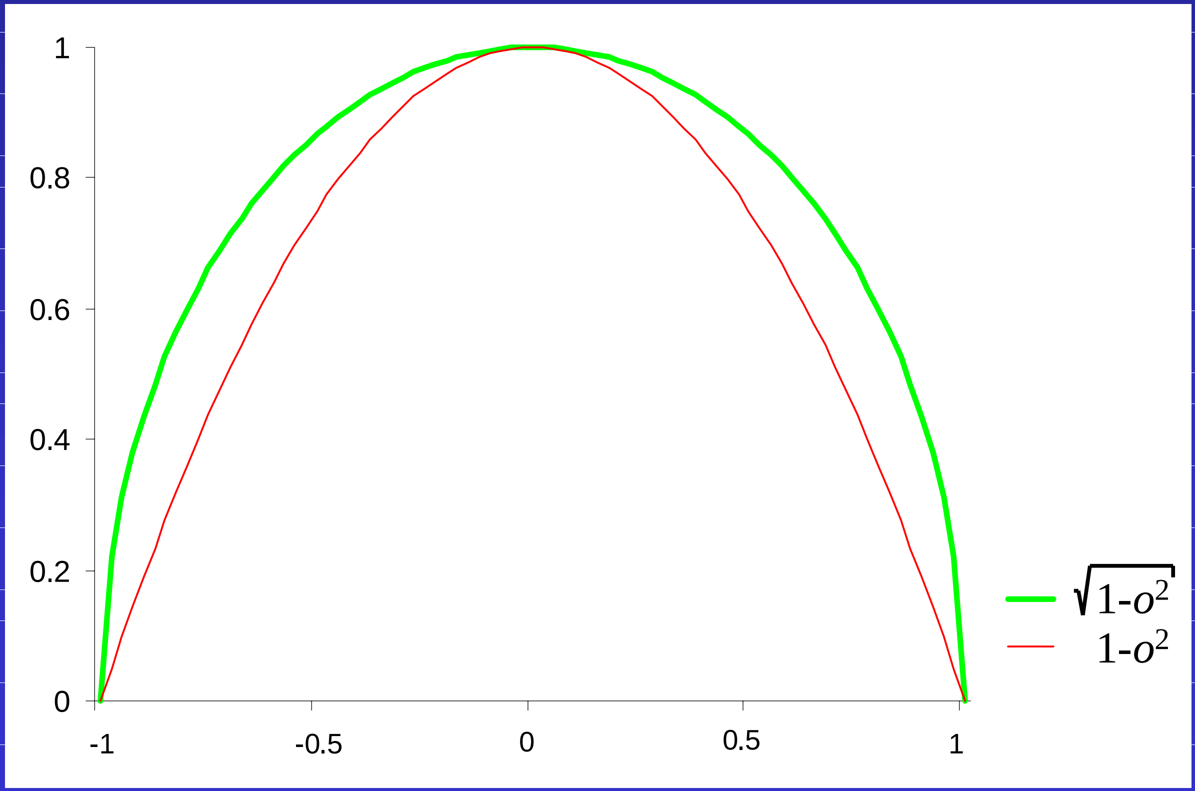
$$\delta_{output\ \ layer} = \left(1 - o^2\right) \cdot sign(t - o) \cdot \left(t - o\right)^2$$

Example:

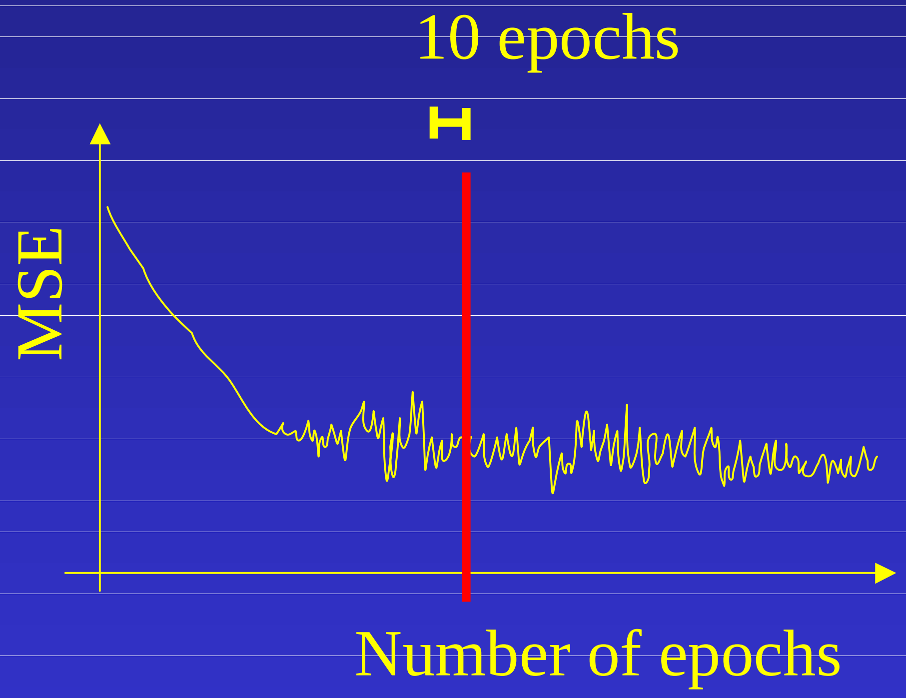
	Pattern 1	Pattern 2	Quotient
$\lvert t - o \rvert$	0.8	0.2	4
$(t - o)^2$	0.64	0.04	16

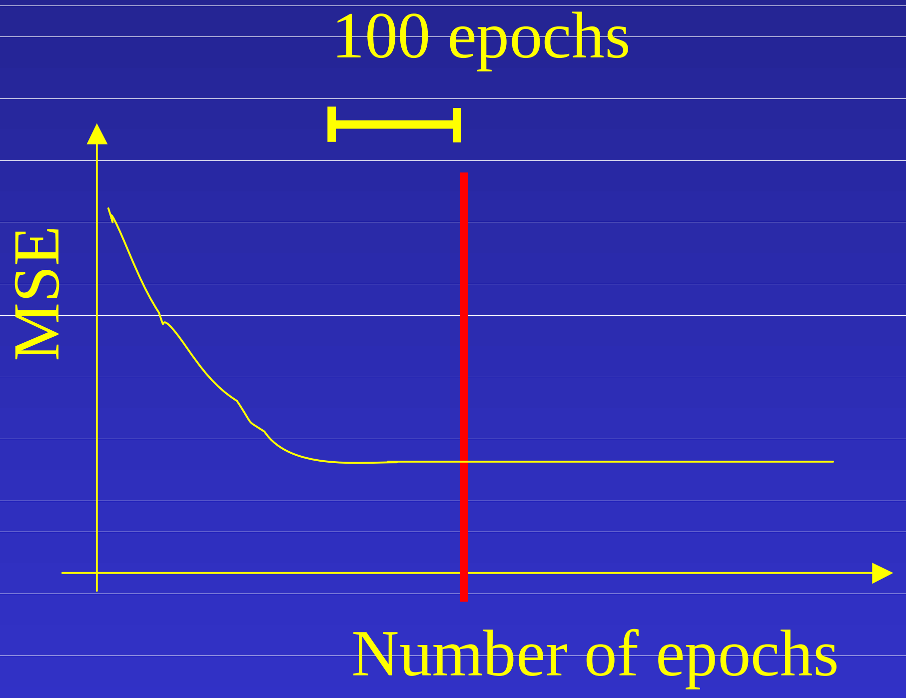
Modifying the Derivation

$$\delta_{output\ \ layer} = \sqrt{1-o^2}\cdot sign(t-o)\cdot(t-o)^2$$

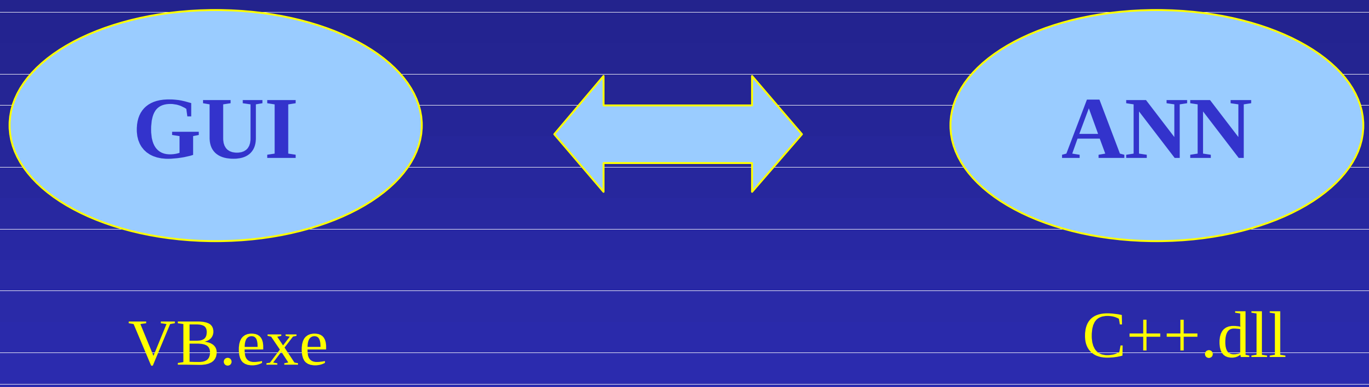


Stop Oscillating and Unsuccessful Learning





The Software



- Compiler settings optimised for speed
- Multithreaded Code
- Extensive use of pointers
- Dynamic arrays
- Avoidance of if-statements or consecutive instructions that hamper pipelining

Approximation of the Exponential Function

Schraudolph, Neural Computation 11, 853-862 (1999):

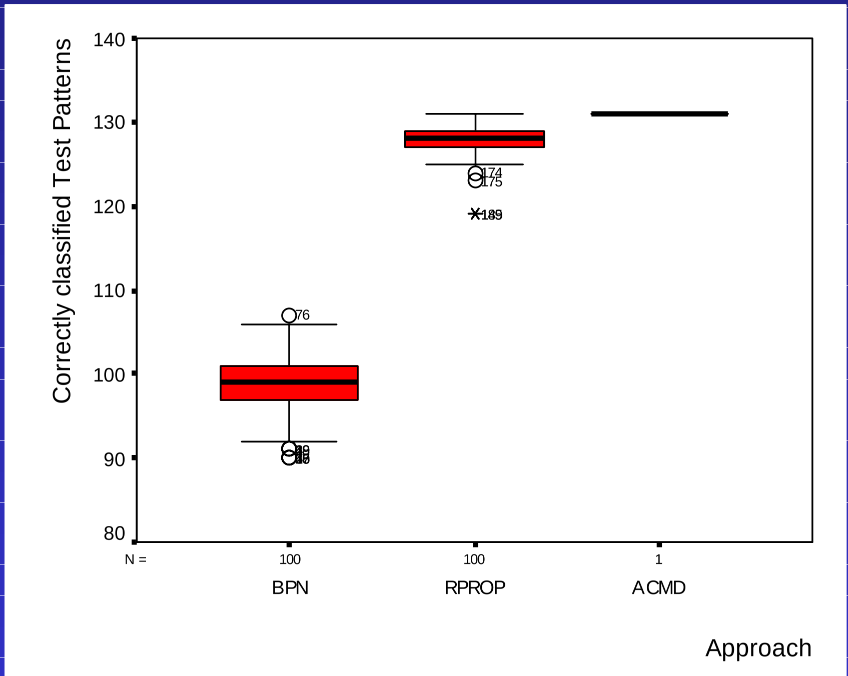
$$\underbrace{\text{sxxxxxxx xxxxmmmm mmmmmmmm mmmmmmmm mmmmmmmm mmmmmmmm mmmmmmmm mmmmmmmm}}_{\text{Float variable: sign (s) + 11-bit exponent (x) + 52-bit mantissa (m)}}$$

Float variable: sign (s) + 11-bit exponent (x) + 52-bit mantissa (m)

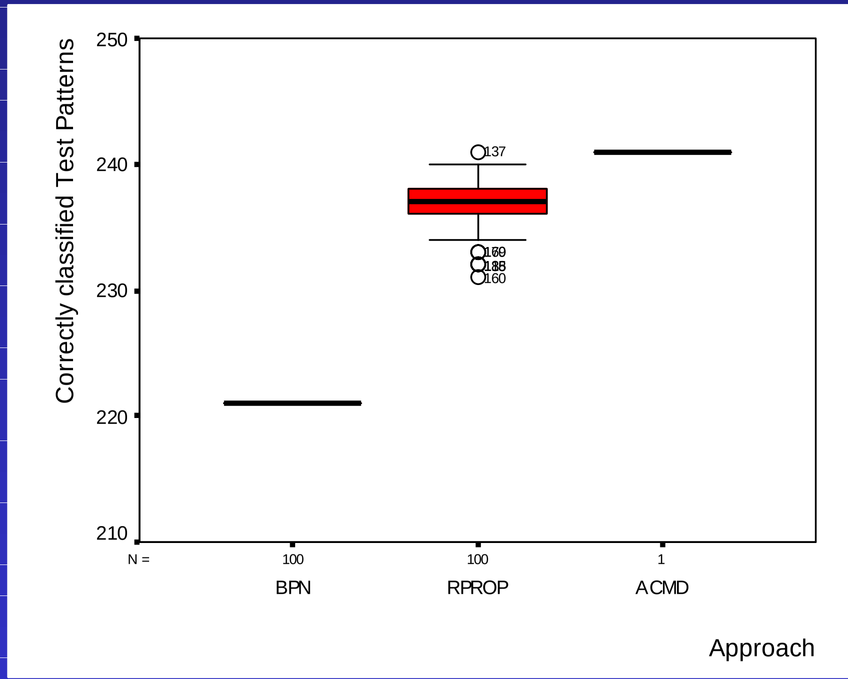
Integer variable A

Integer variable B

Results



Edible Oil



Honey

Conclusion

If you decide to use an ANN for pattern recognition of EN signals or whatever biosignals – it is worthwhile to **consider suitable ANN strategies.**

Thank you!

linder@medinf.mu-luebeck.de

Tel.: ++49 451 500 6633