

Metody analizy danych – zastosowanie sztucznych sieci neuronowych i nie tylko.

Piotr Płoński

Politechnika Warszawska, WEiTI

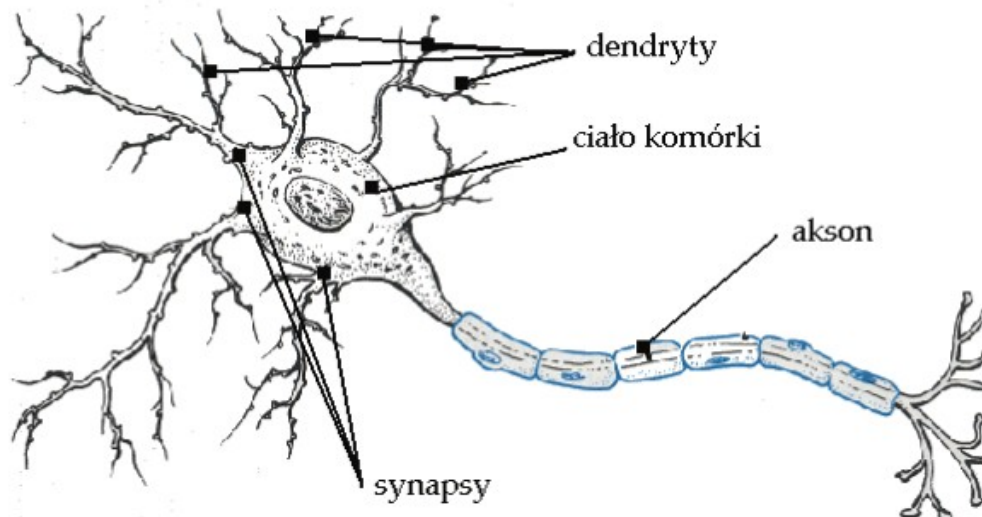
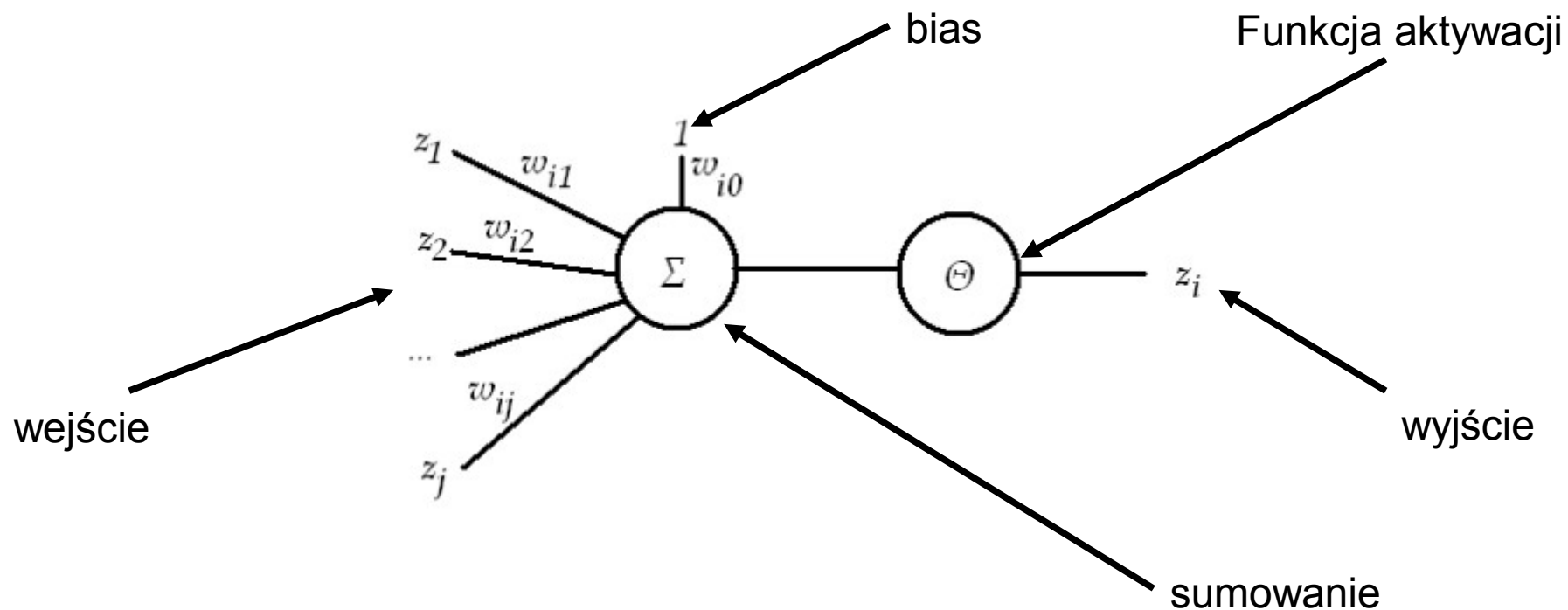
Plan prezentacji

1) MLP, SOM, kNN

- Opis architektury
- Algorytm uczenia
- Przykład działania
- Zastosowanie

2) Klasyfikacja różno-wymiarowa

Perceptron Wielowarstwowy Multi Layer Perceptron (MLP)



Funkcja aktywacji

- funkcja liniowa:

$$y_k = g(a_k) = \beta a_k,$$

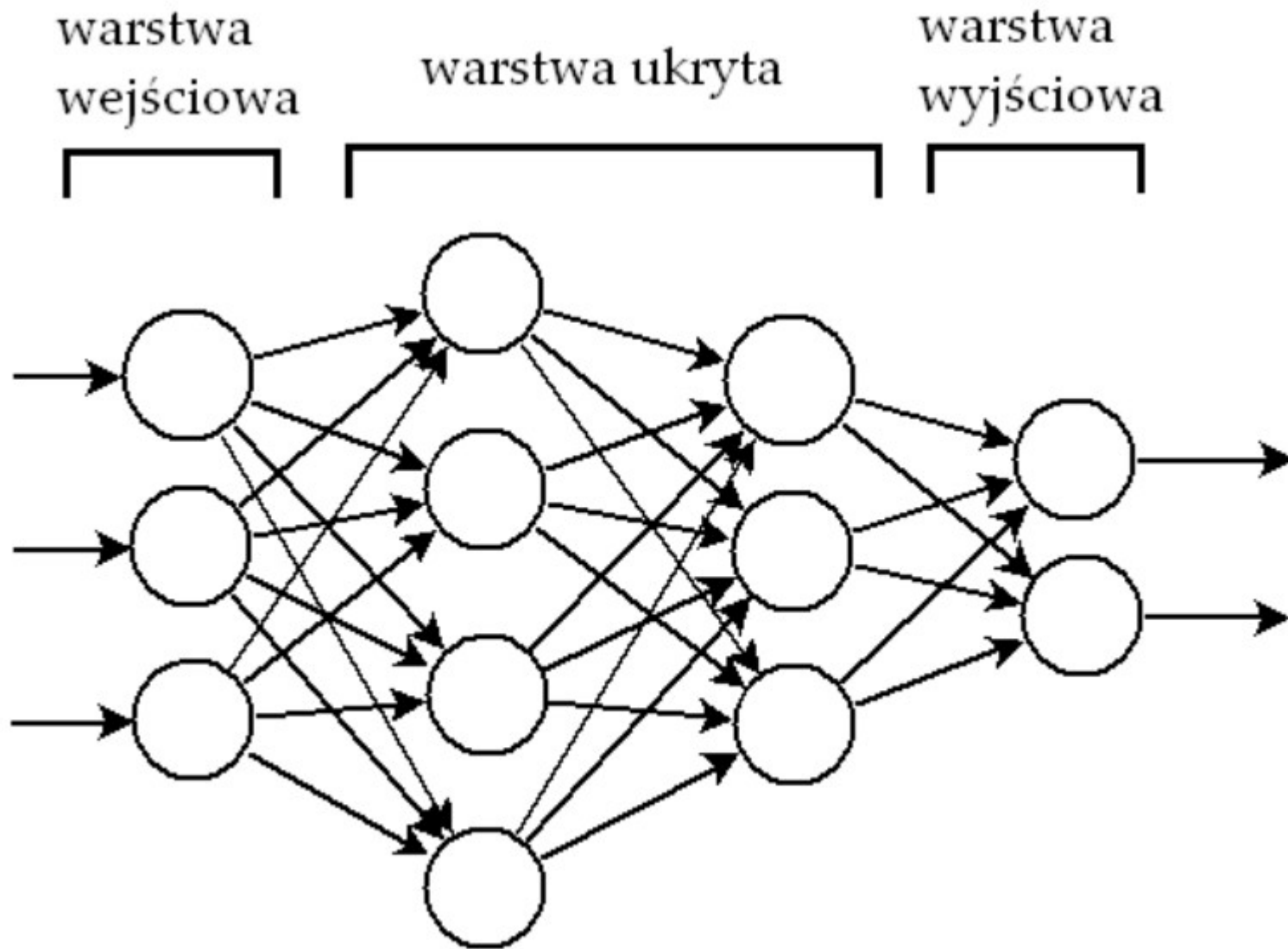
- funkcja sigmoidalna unipolarna:

$$y_k = g(a_k) = \frac{1}{1 + \exp(-\beta a_k)},$$

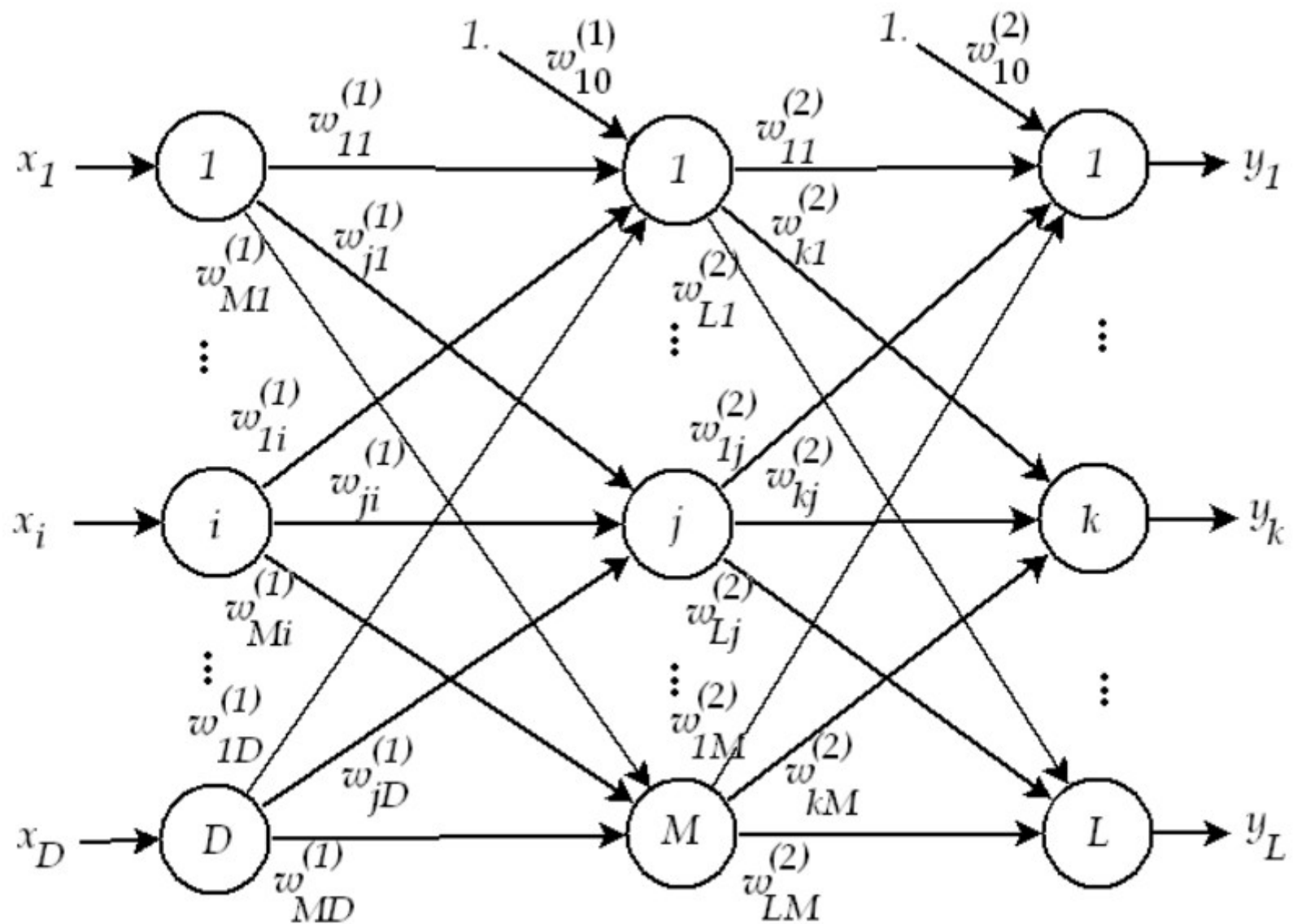
- funkcja sigmoidalna bipolarna:

$$y_k = g(a_k) = \frac{2}{1 + \exp(-\beta a_k)} - 1.$$

Wiele warstw  sieć



Propagacja sygnału



Dobór wag – uczenie sieci

$$E(\mathbf{W}) = \sum_N \sum_K (y_k^{(n)} - t_k^{(n)})^2$$
$$\chi^2 = \sum_N \sum_K \frac{1}{\sigma_n^2} (y_k^{(n)} - t_k^{(n)})^2$$

FUNKCJE BŁĘDU

wartość wyjściowa wartość żądana

ZMIANA WAG

$$\Delta \mathbf{W} = -\eta \nabla E$$

GRADIENT FUNKCJI BŁĘDU

$$\nabla E = \left[\frac{\partial E}{\partial w_1}, \frac{\partial E}{\partial w_2}, \dots, \frac{\partial E}{\partial w_n} \right]$$

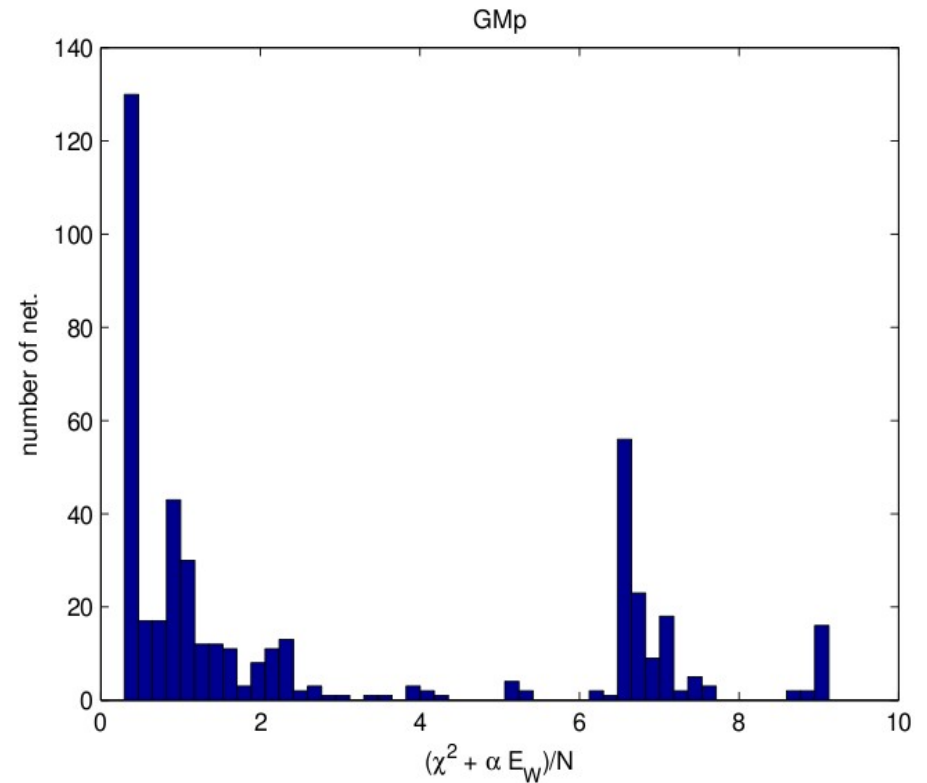
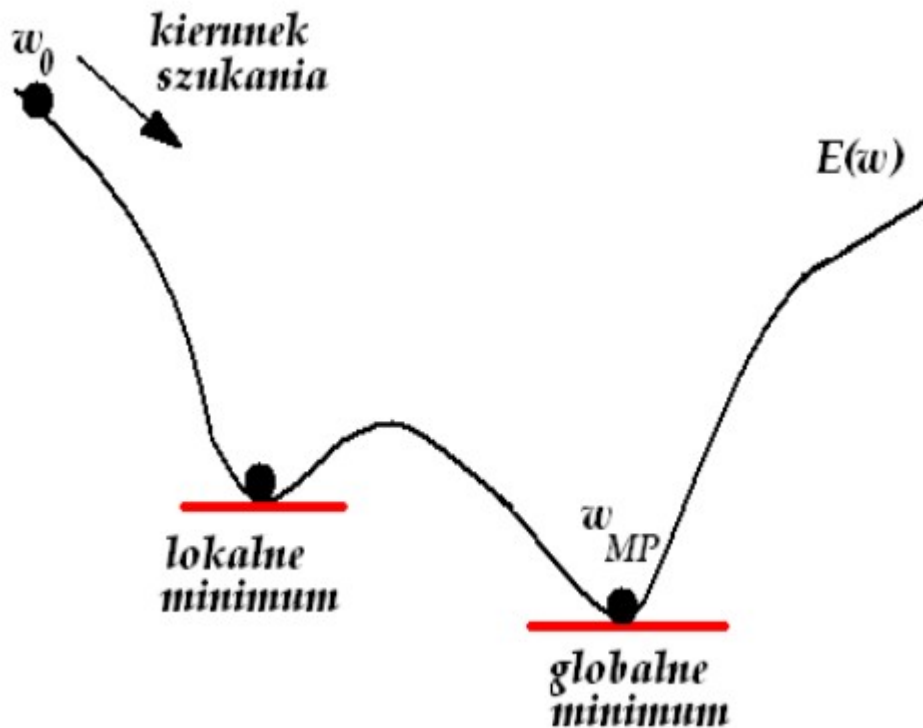
UAKTUALNIENIE WAG

$$\mathbf{W}(k+1) = \mathbf{W}(k) + \Delta \mathbf{W}$$

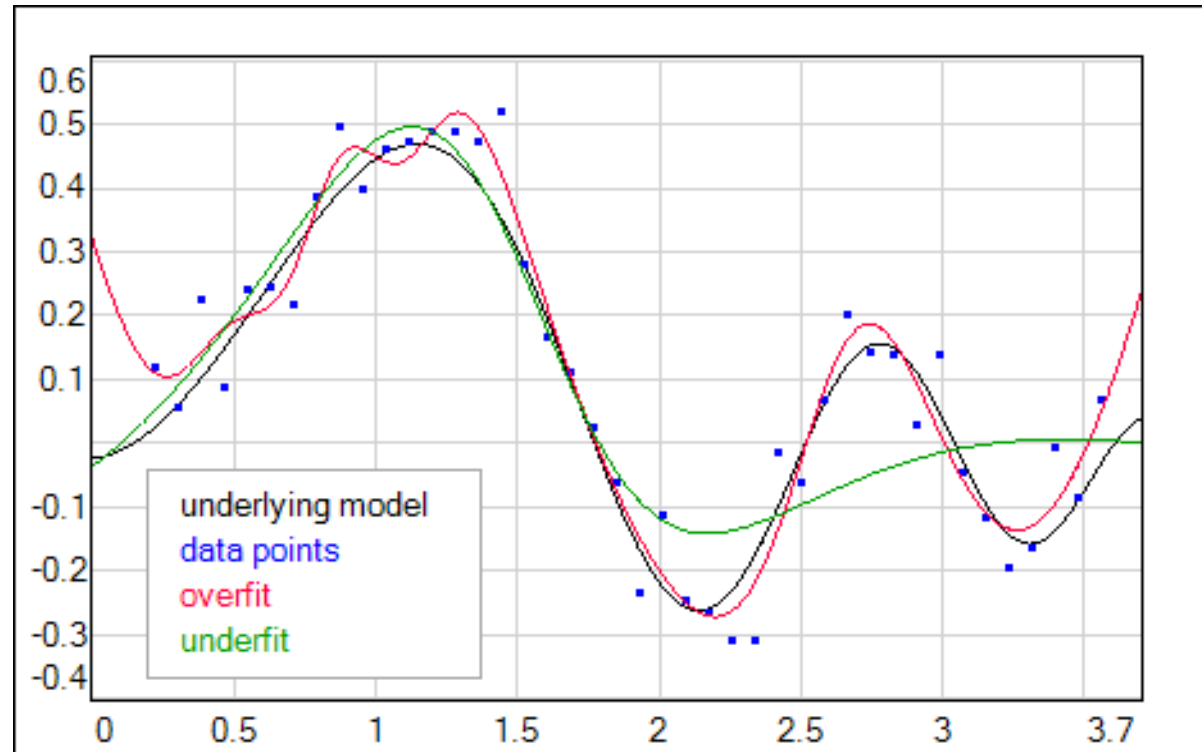
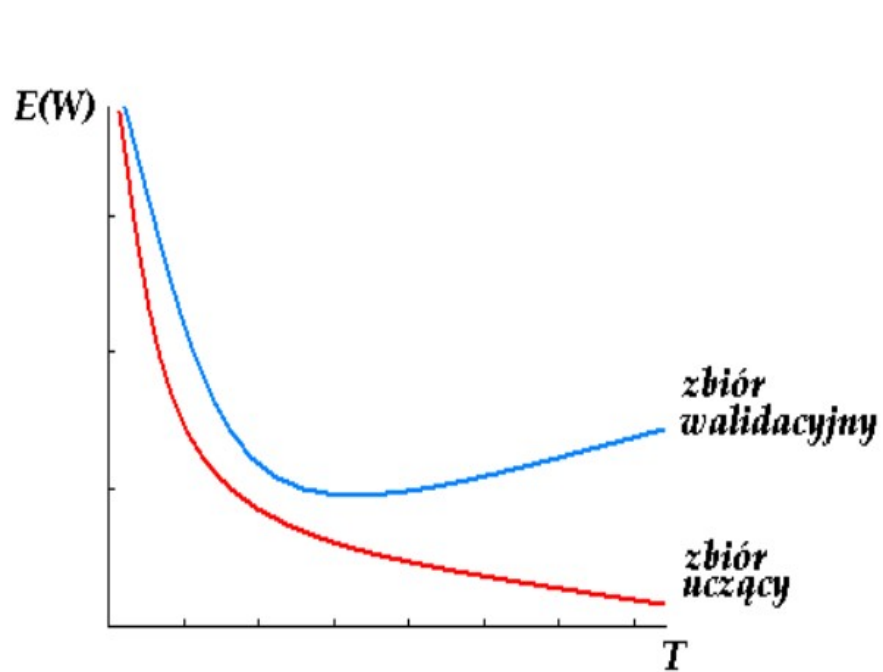
Algorytmy uczące

- algorytm propagacji wstecznej
- quickprop
- gradientów sprzężonych
- Levenberga-Marquardta

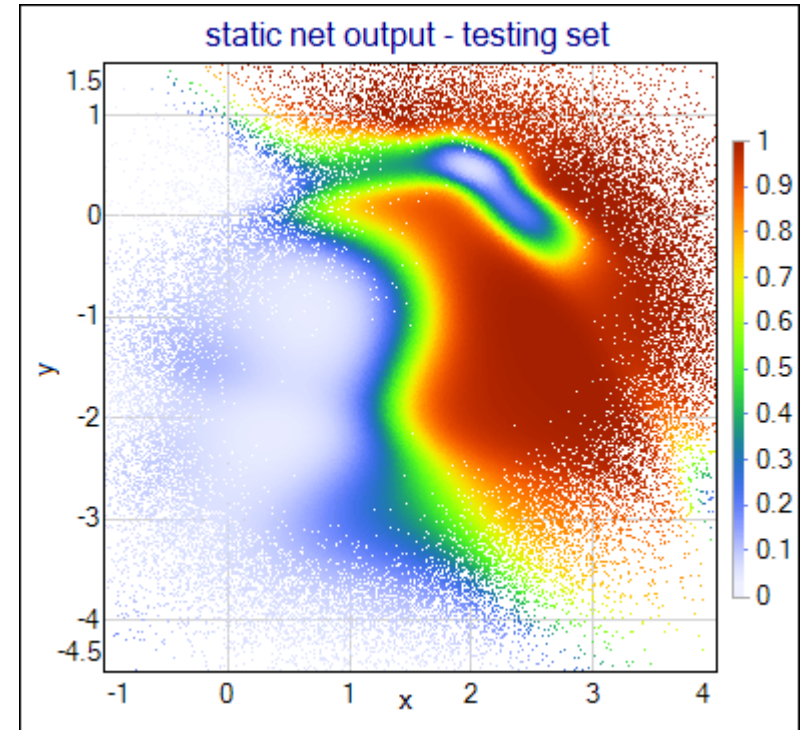
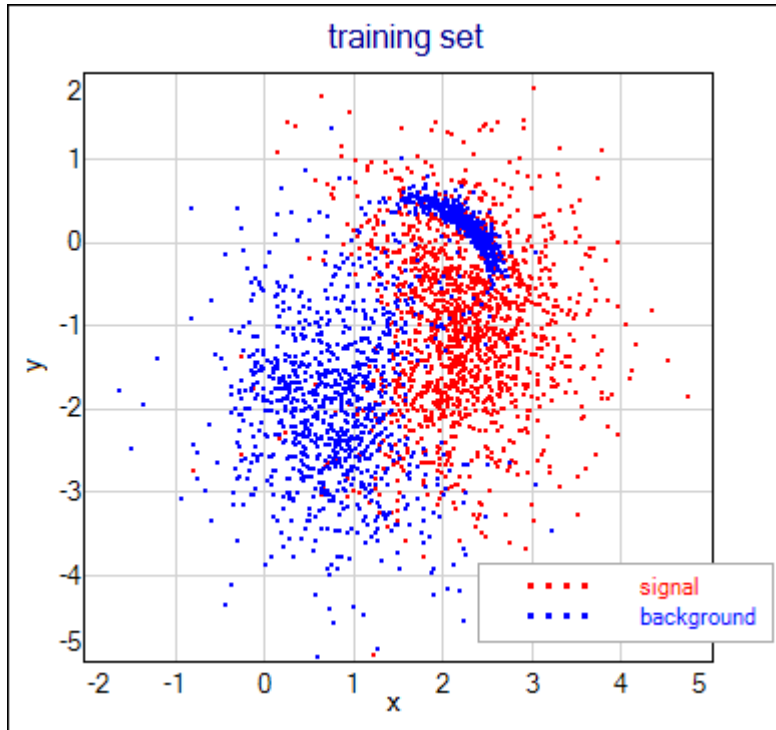
Problemy w uczeniu – lokalne minima



Problemy w uczeniu overfit i dobór struktury



Przykład - klasyfikacja



PRÓBKA	x	y	KLASA	
			1	2
1	4,32	13,76	1	0
2	54,3	0,11	0	1
3	43,49	1,22	0	1
4	52,47	2,53	?	?

Klasyfikacja w fizyce

- * Successful discrimination of gluon from quark jets have been reported in both $e+e-$ annihilation and hadron induced reactions.
- * Identifying b-quarks in annihilation using four momenta of the leading hadrons.

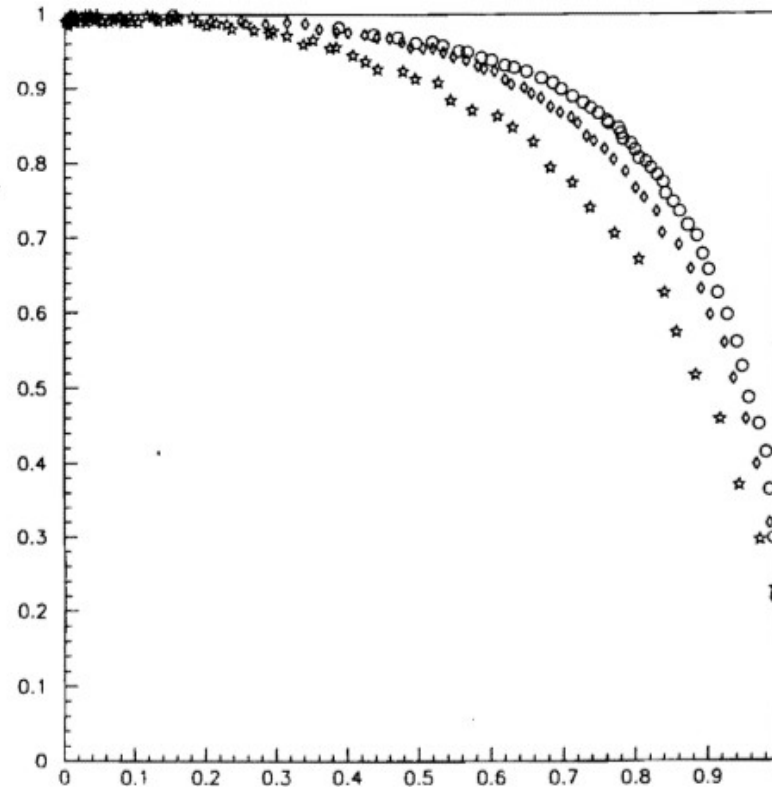


Figure 2: Purity/efficiency for b-quark discrimination (from ref. [20]) using single variable discriminator (star) and multivariate (lozenge) and ANN (open circle) respectively.

Klasyfikacja

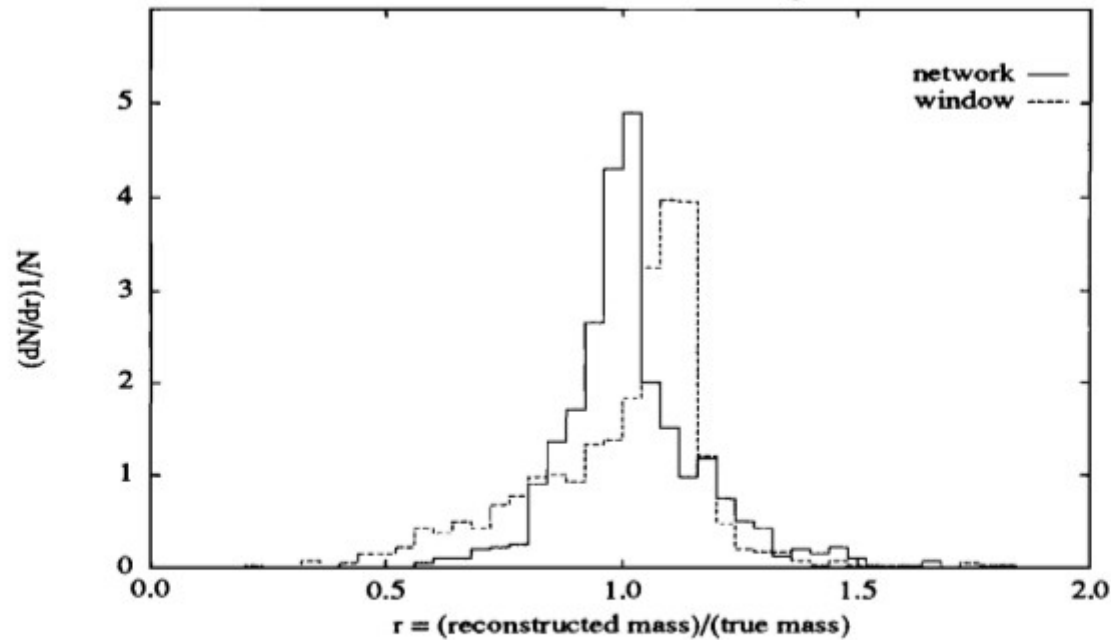
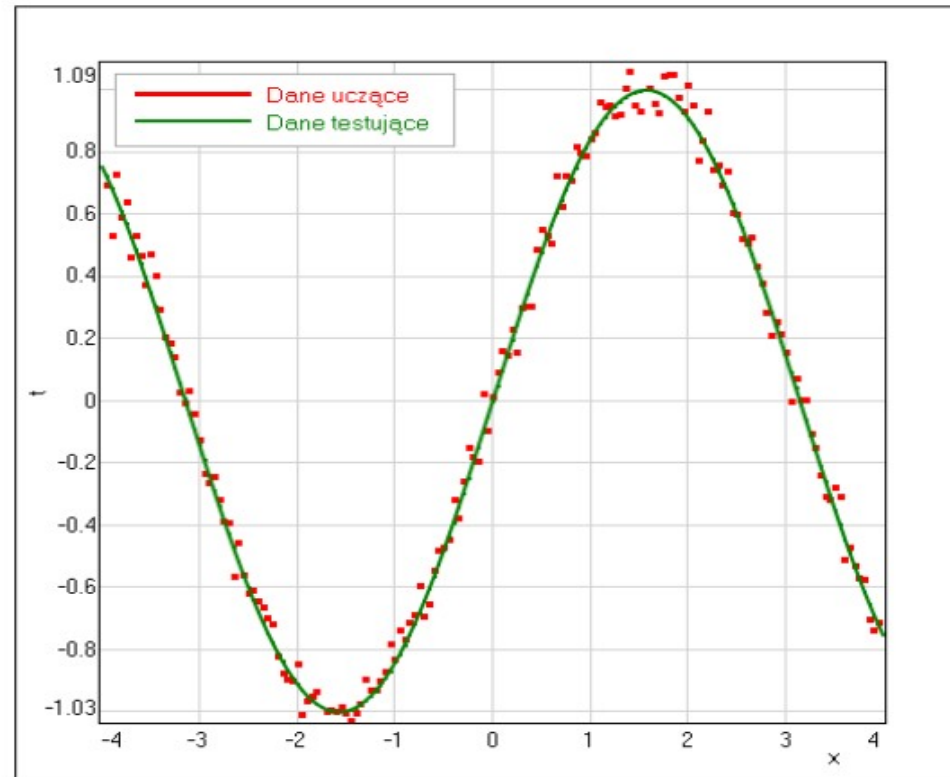


Figure 3: The reconstructed mass ($M_{W,Z}$) divided by the true mass ($M_{W,Z}^0$) using the neural network method (full line) and the conventional “window” method (dashed line) with $R = 0.8$.

Przykład - aproksymacja



PRÓBKA	X	Y
1	13,76	13,76
2	0,11	0,11
3	1,22	1,22
4	2,53	?

Aproksymacja

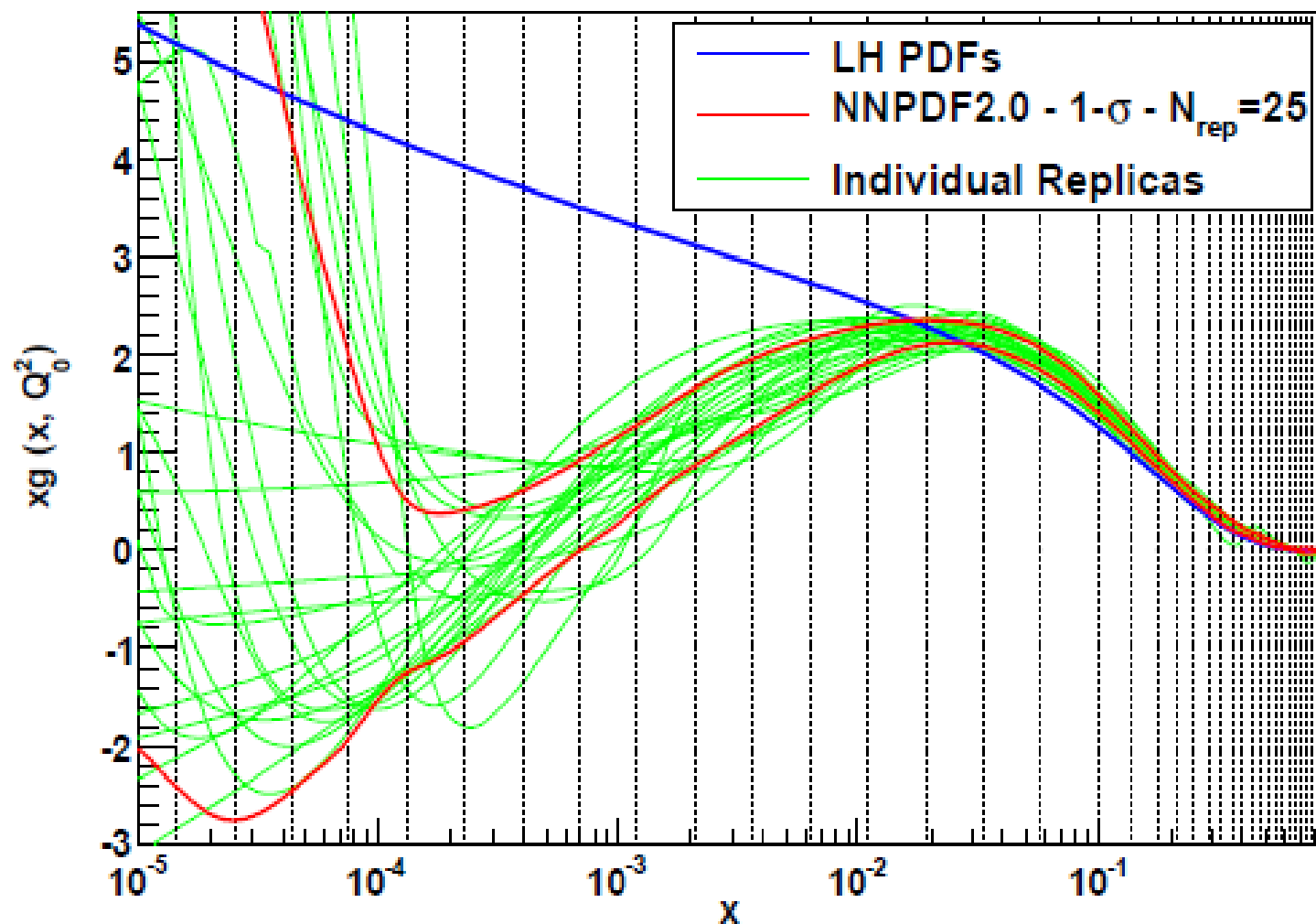


Figure 4: The sampling grid of x values used for evolution superposed to the Les Houches gluon and 25 replicas of the gluon distribution from the NNPDF2.0 set at the starting scale.

Aproksymacja

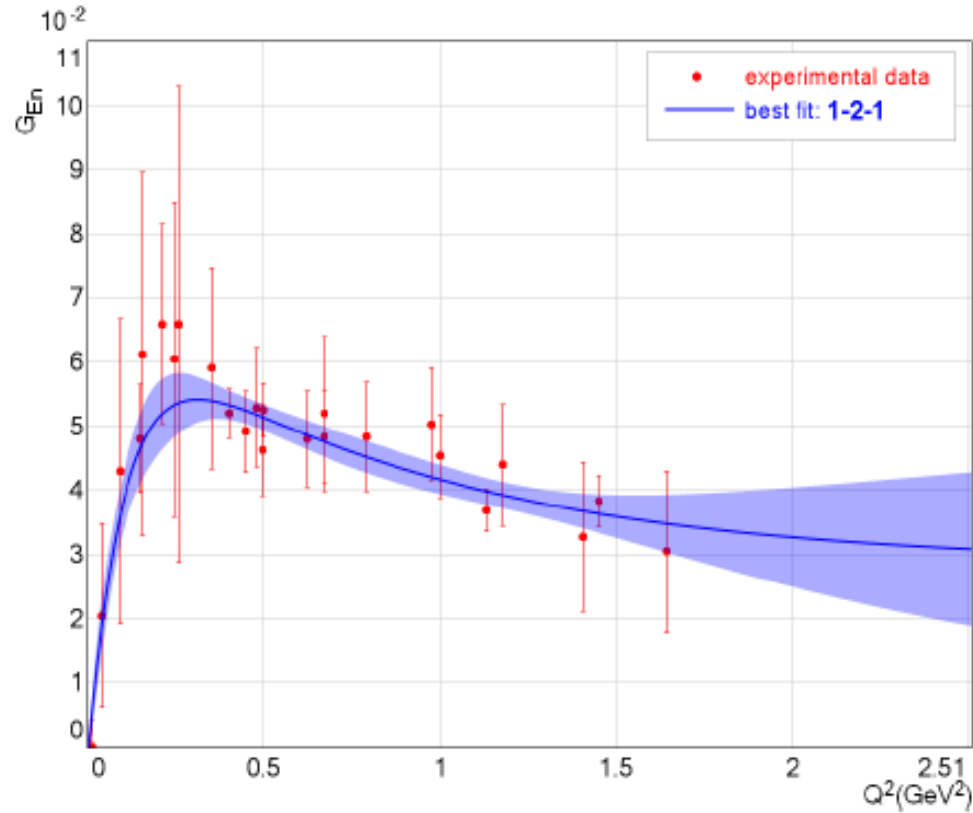
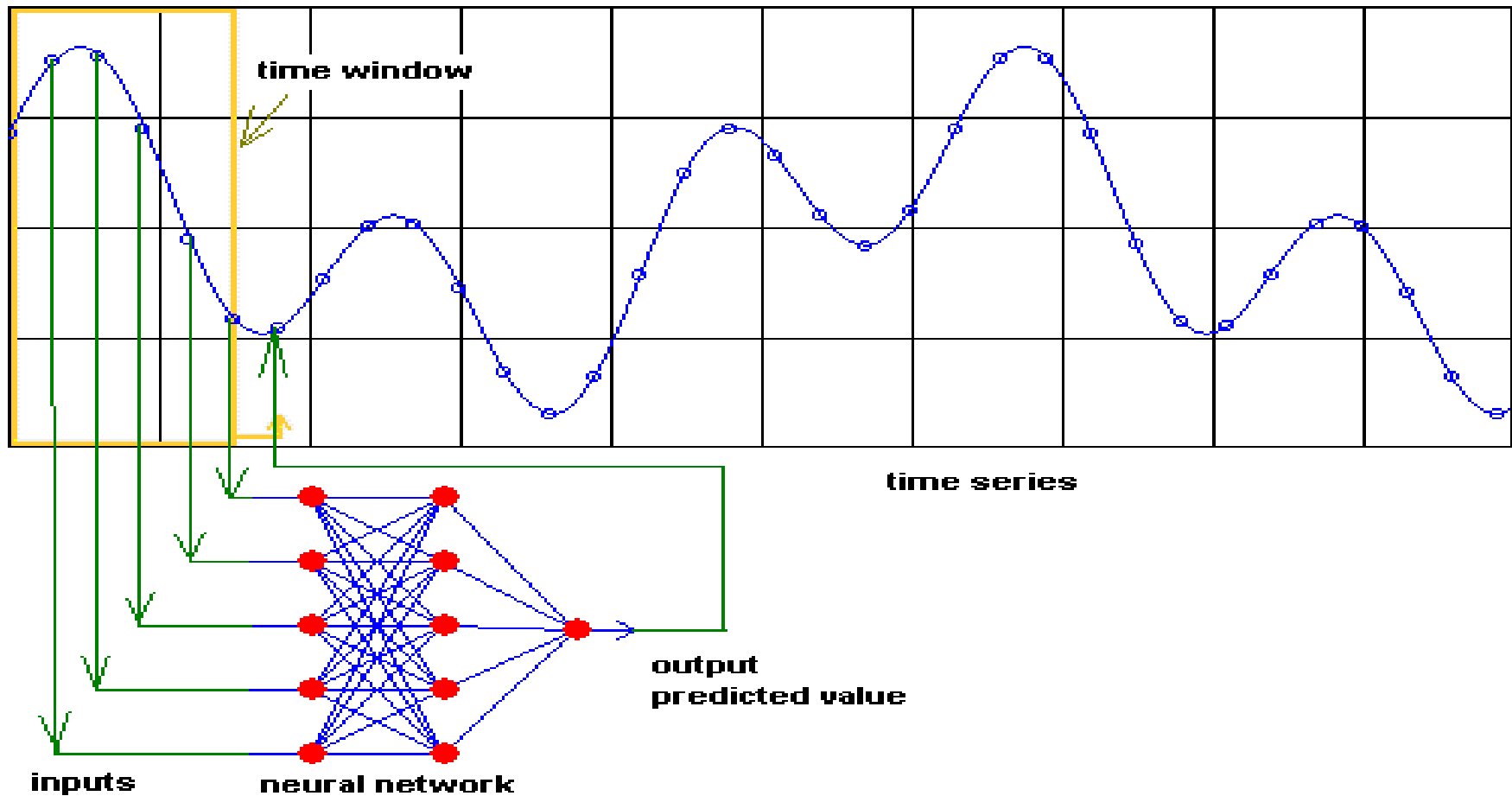


Figure 16: The best fit of G_{En} data given by the 1-2-1 network. The blue area denotes fit uncertainty computed with Eq. 3.16. The experimental data is the same as the one discussed in Ref. [30].

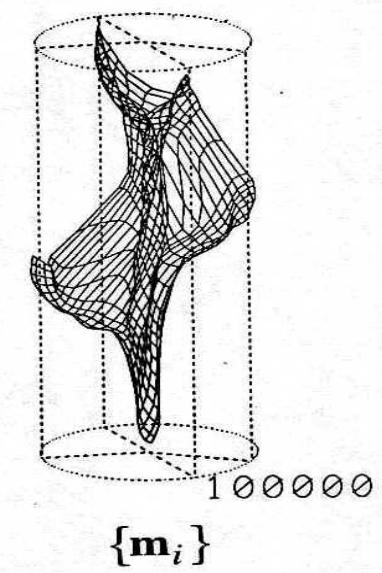
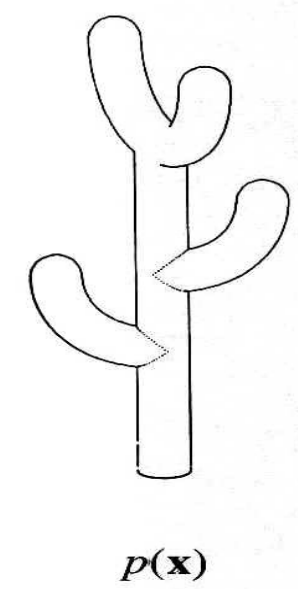
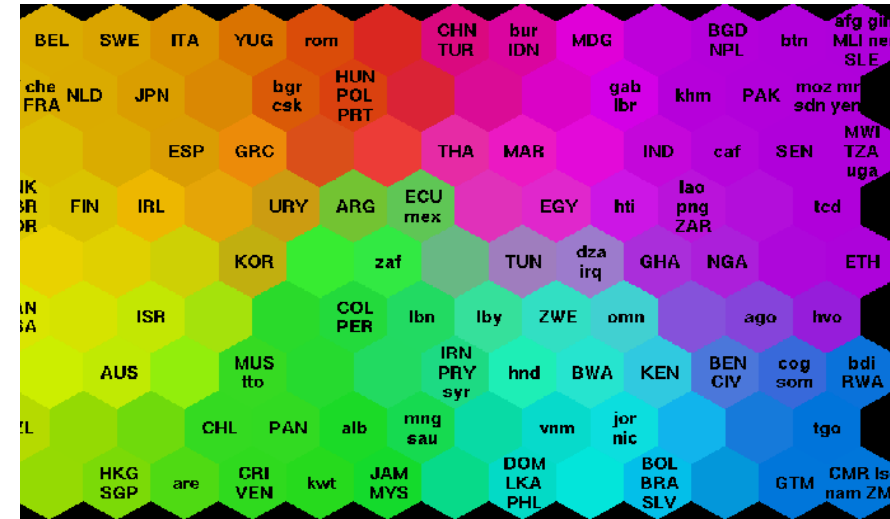
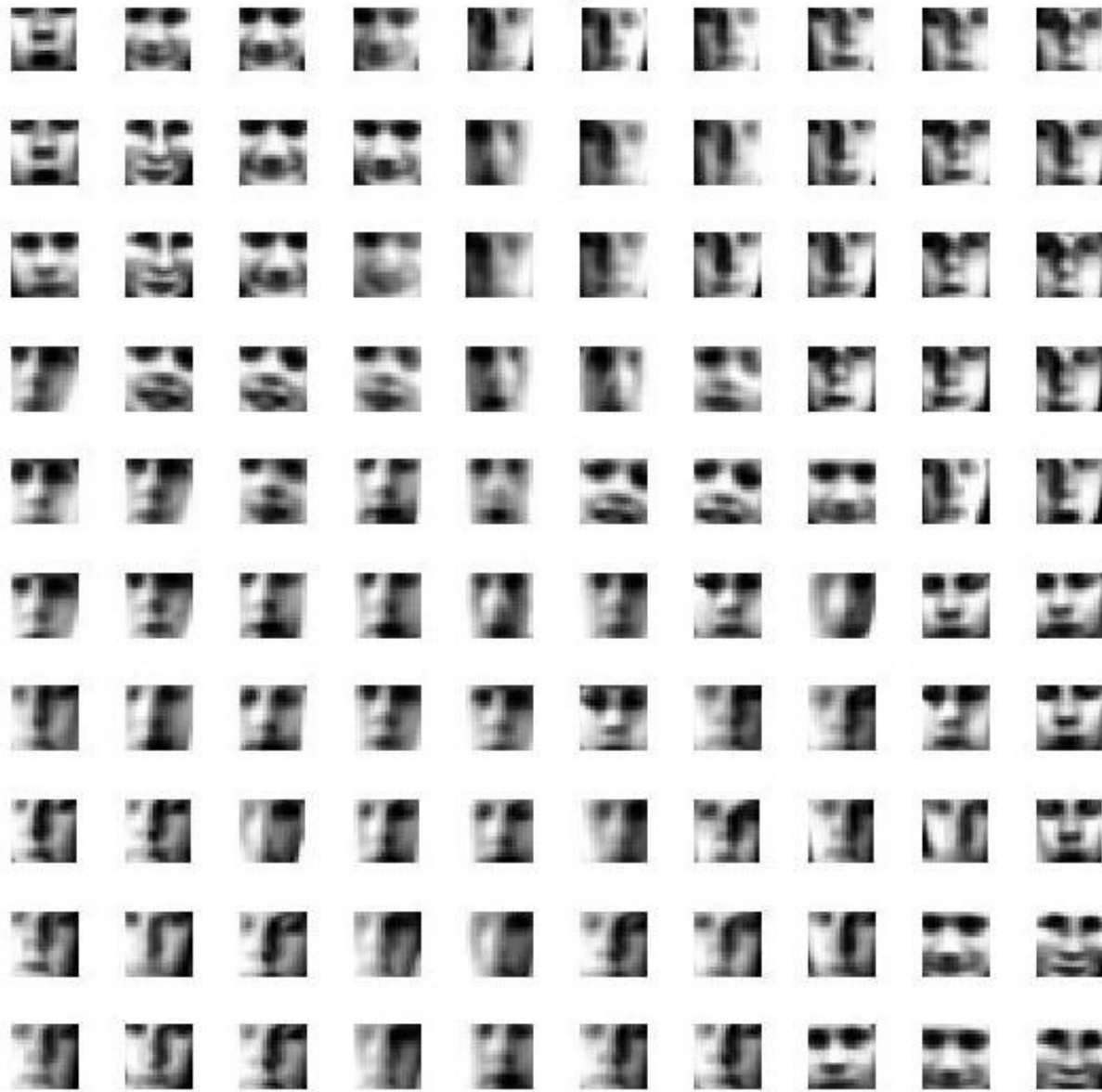
Przykład - predykcja



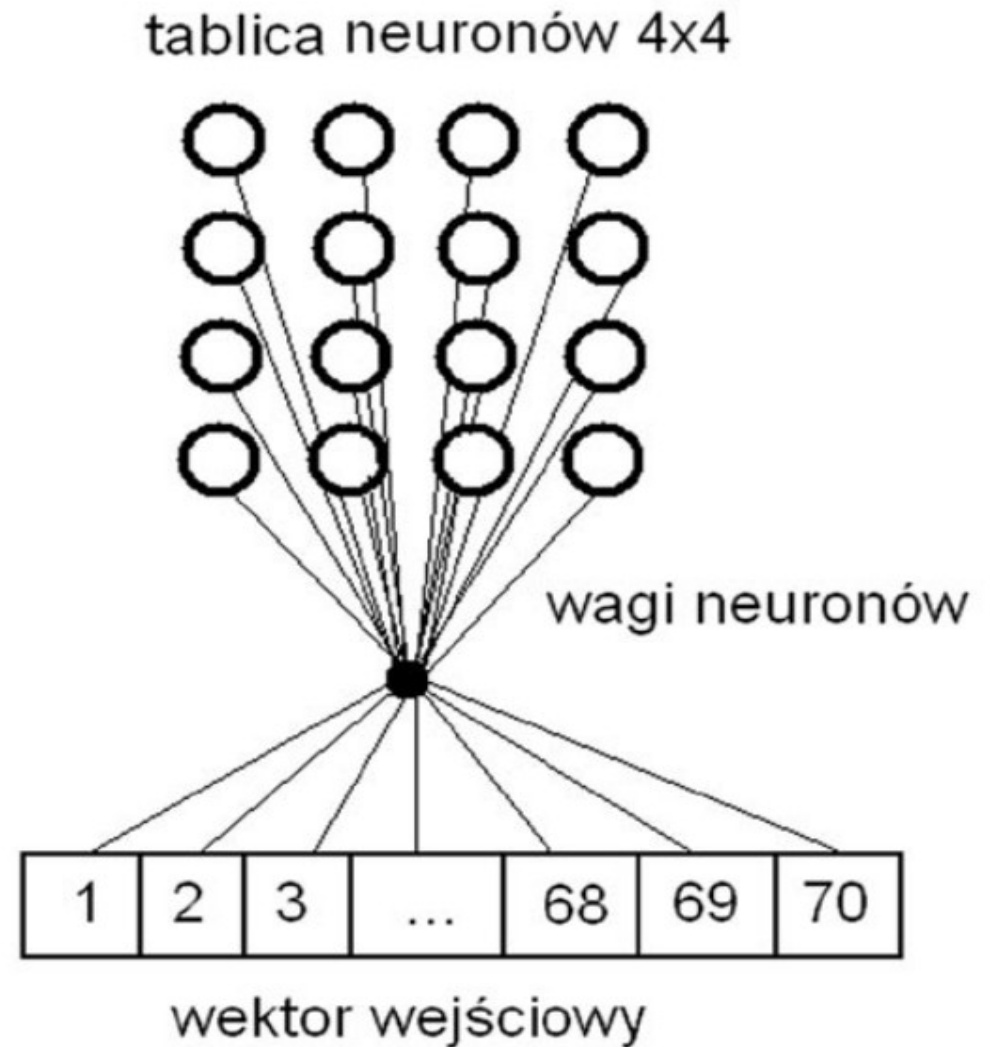
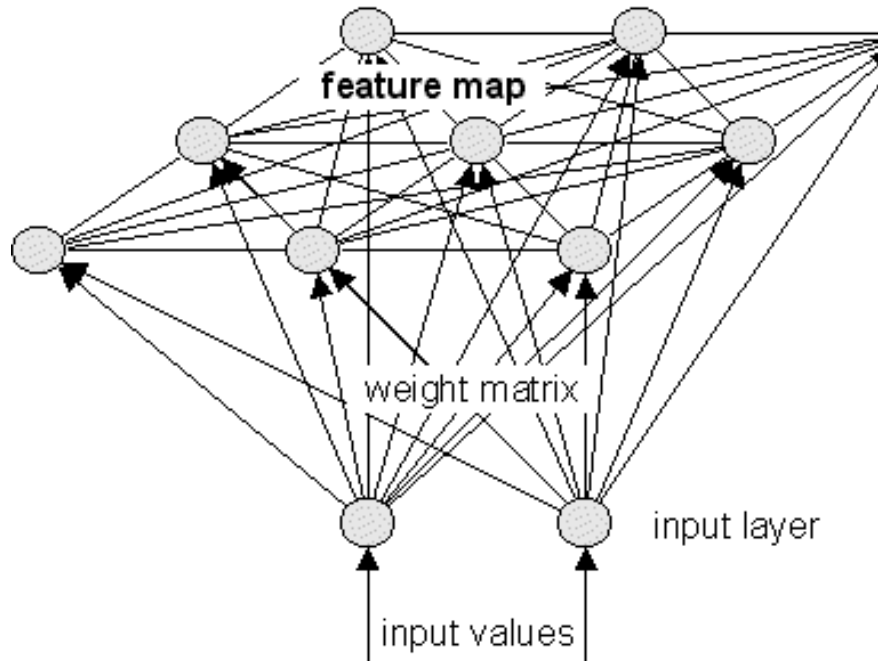
PRÓBKA	CZAS	X	Y
1	1	13,76	13,76
2	2	0,11	0,11
3	3	1,22	1,22
4	4	2,53	?

Sieci samoorganizujące się sieci Kohonena

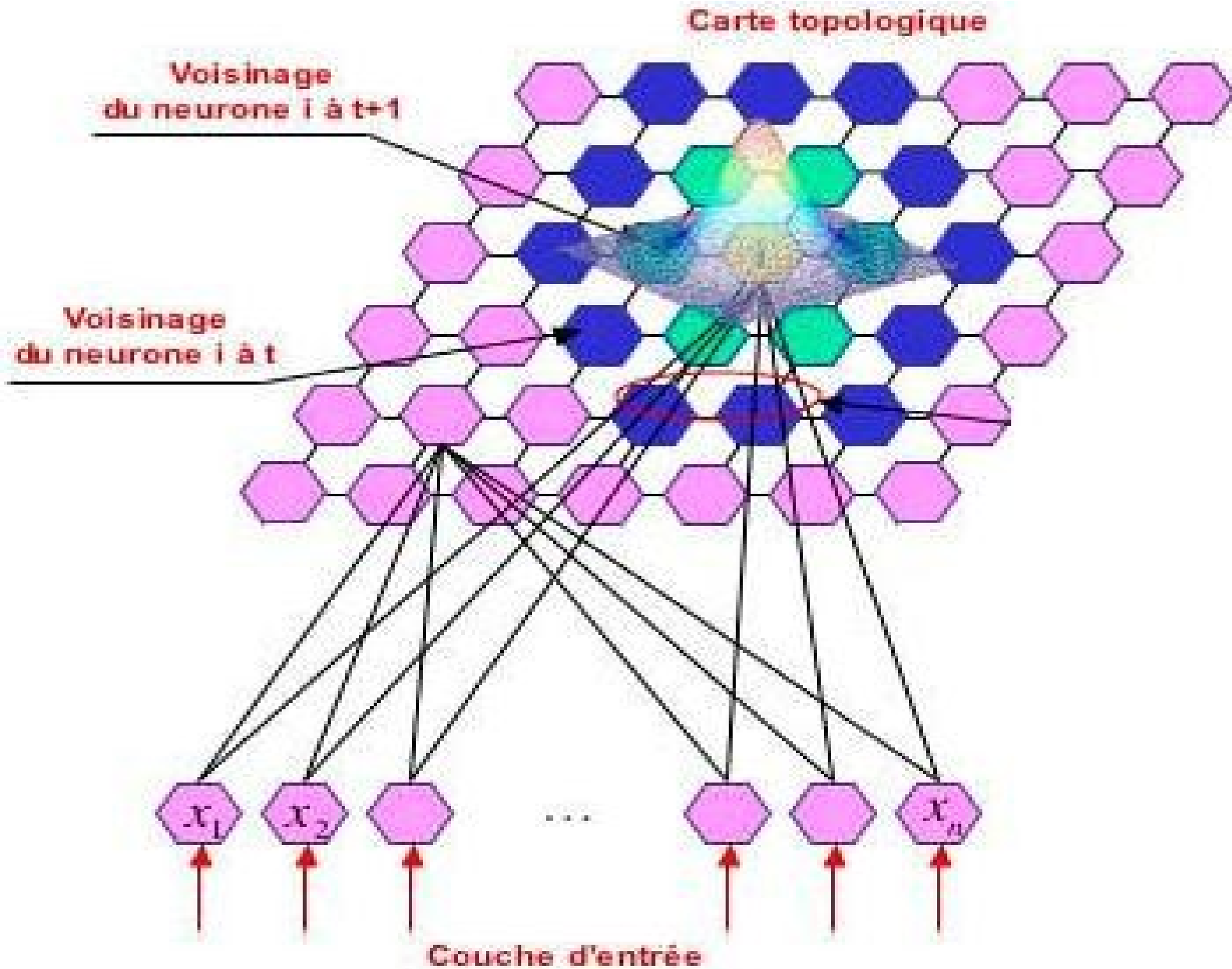
10x10 SOM clustering results



Self Organising Map - Architektura



SOM - sąsiedztwo



Uczenie – reguła Hebba

$$D_{ij}(W, X) = \sqrt{\sum_{q=1}^N (W_q - X_q)^2}$$

← odległość między neuronem a wektorem wejściowym

AKTUALIZACJA WAG

$$W(t+1) = W(t) + L(e) \cdot S(i, j) \cdot (X(t) - W(t))$$

$$L(e) = L_0 \exp\left(-\frac{e}{\lambda}\right)$$

współczynnik
uczenia

$$S(i, j) = \exp\left(\eta \sqrt{(B_i - i)^2 + (B_j - j)^2}\right)$$

FUNKCJA
SĄSIEDZTWA

Best Matching Unit

Klastrowanie - przykład

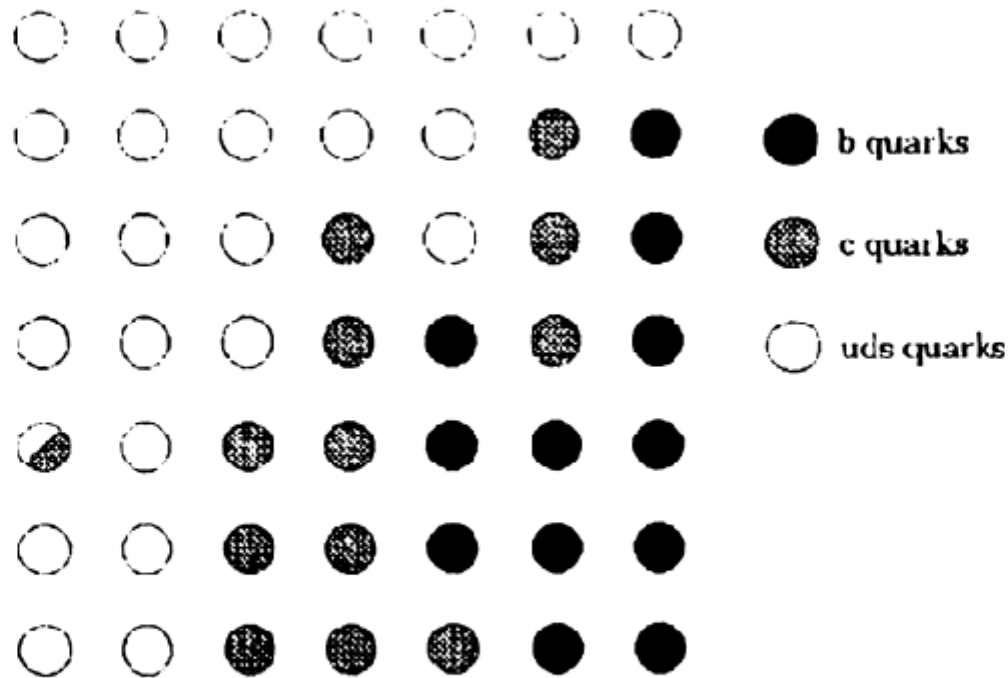
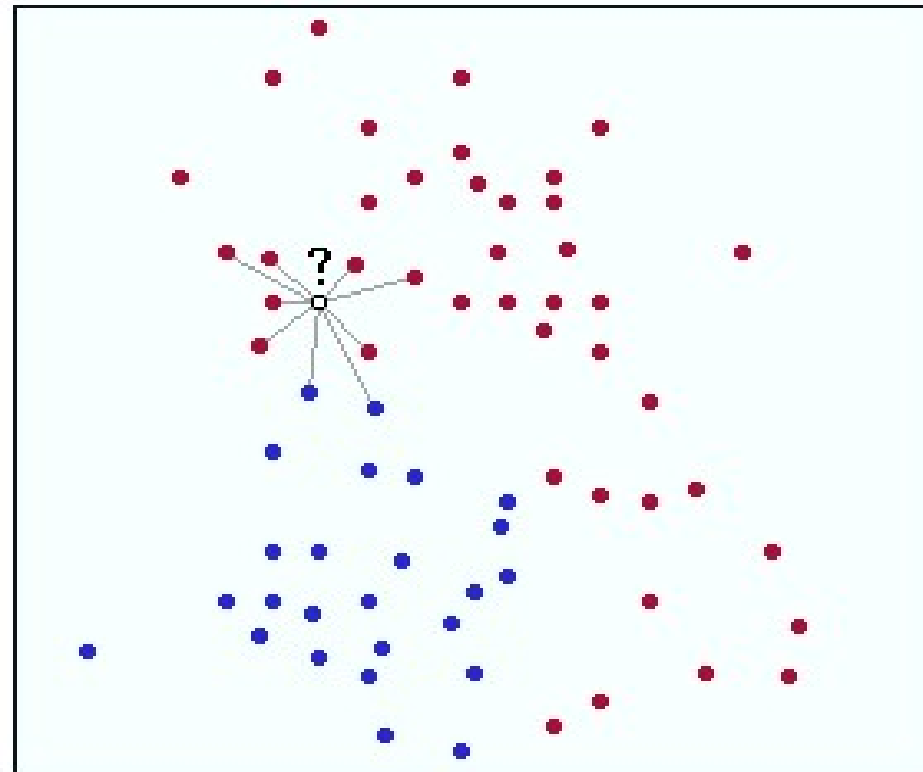


Figure 7: The resulting map for uds- c- and b-quarks. The shading indicates the dominant flavor for the units. The units are numbered as in fig. 6.

k -NN = algorytm k najbliższych sąsiadów

Algorytm:

1. Wyznacz odległość między sprawdzanym punktem P a wszystkimi pozostałymi.
2. Wybierz k (ustalona z góry liczba) najbliższych punktów do P .
3. Wynikiem jest średnia wartość z k najbliższych sąsiadów.



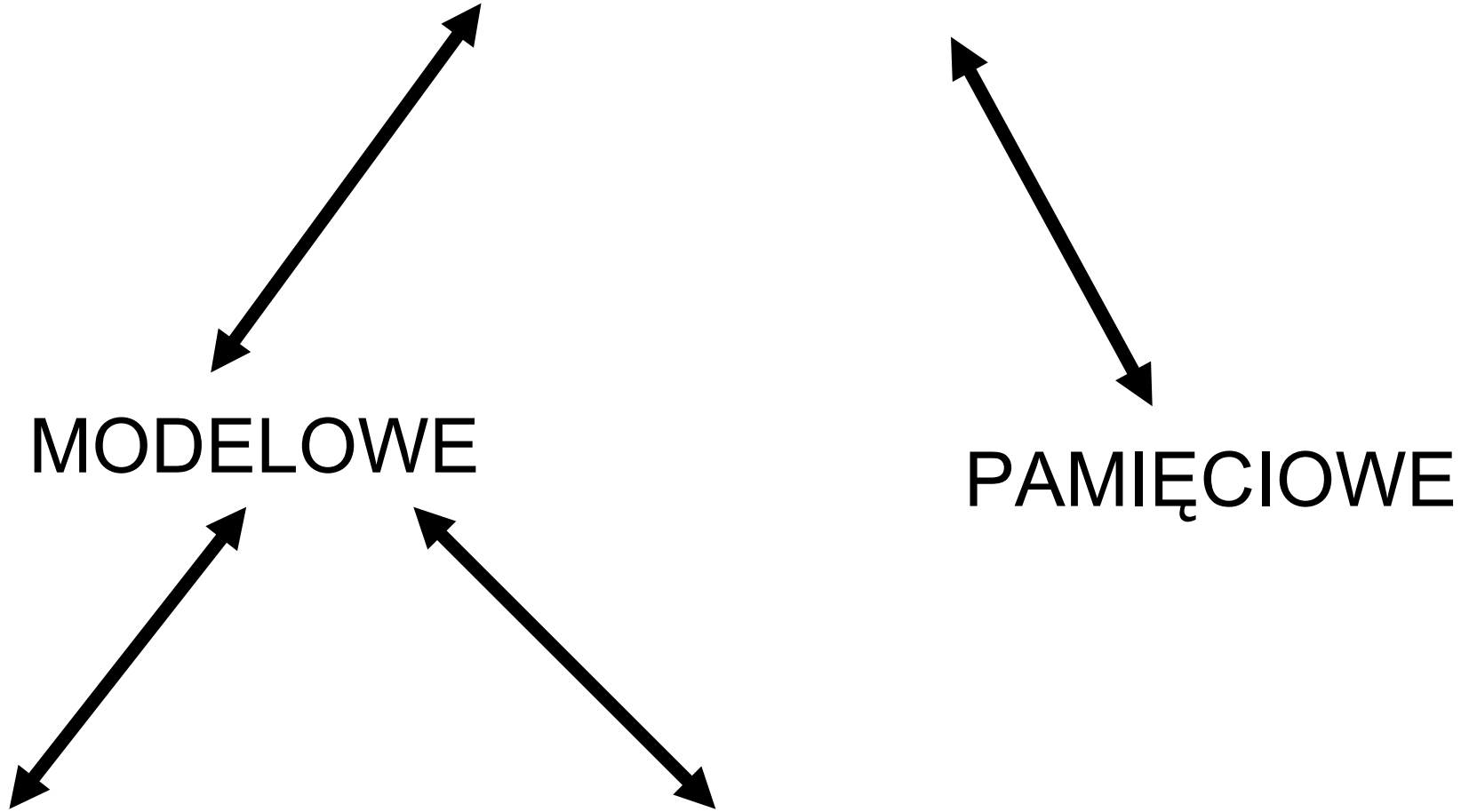
Podział metod

MODELOWE

PAMIĘCIOWE

Uczenie z
nauczycielem

Uczenie bez
nadzoru

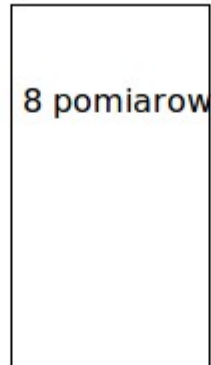


Klasyfikacja danych różno-wymiarowych

zdarzenie 1



zdarzenie 2



zdarzenie 3

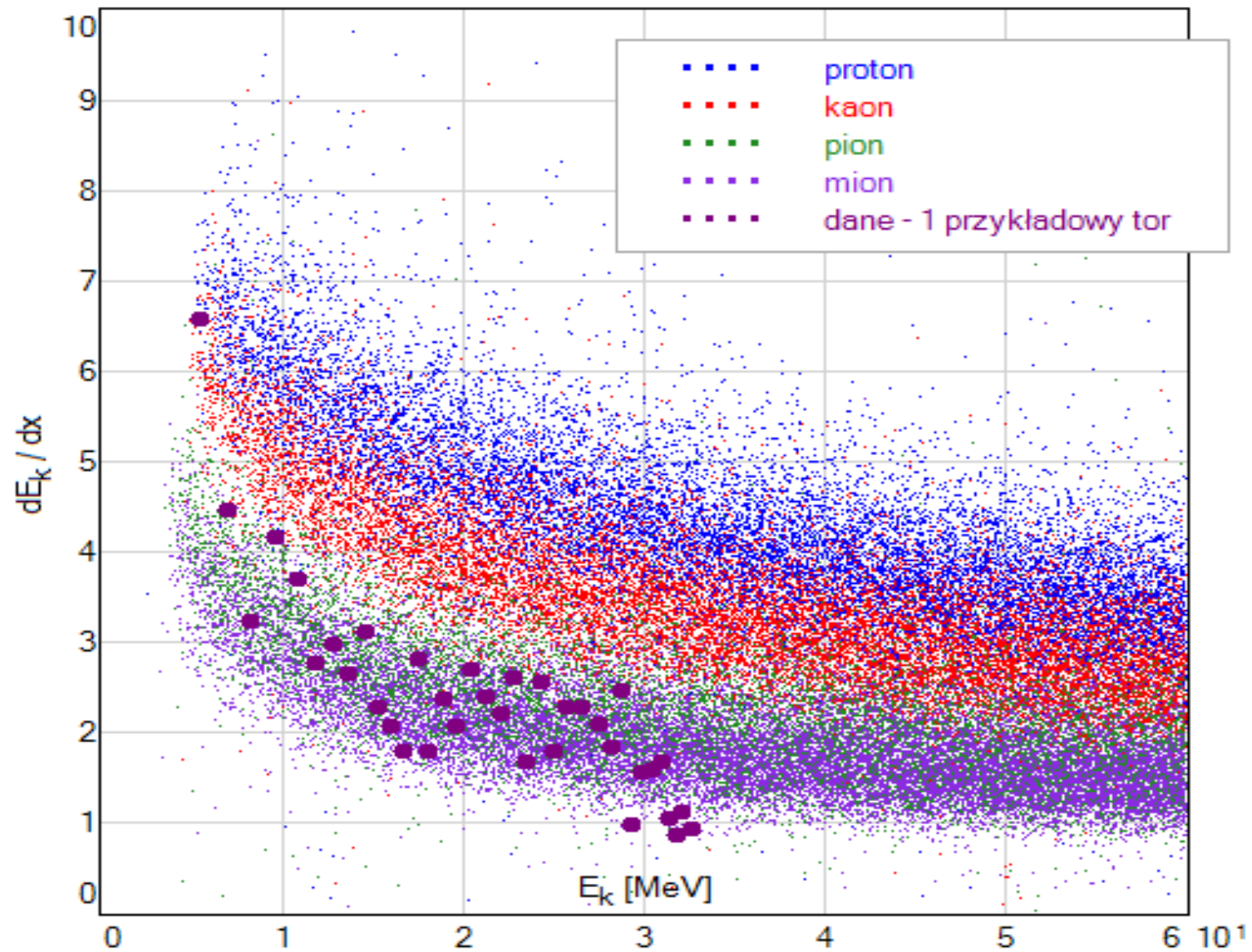


Stare podejście:

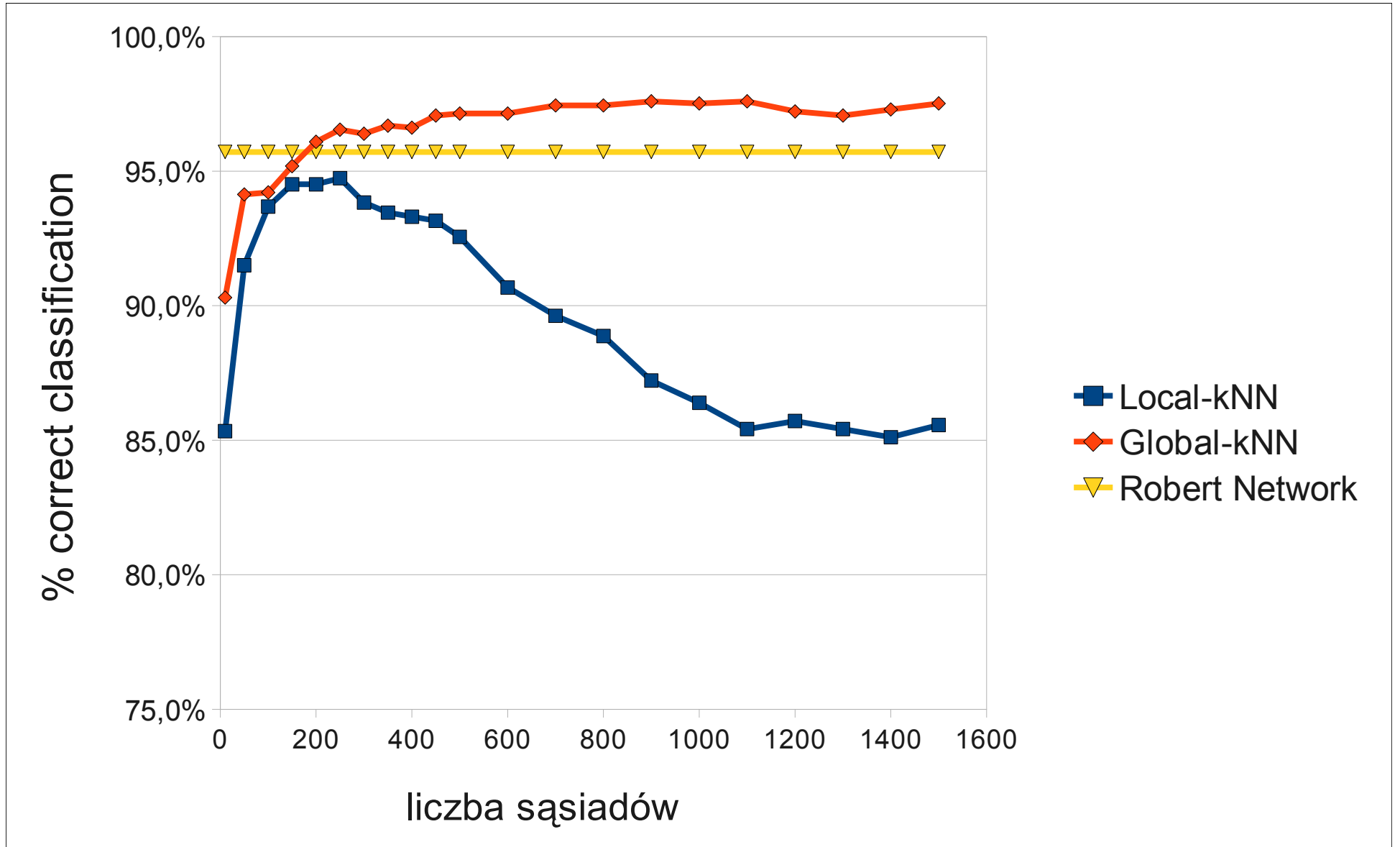
- uśrednianie do stałego wymiaru
- lub klasyfikacja pojedynczych pomiarów i uśrednianie

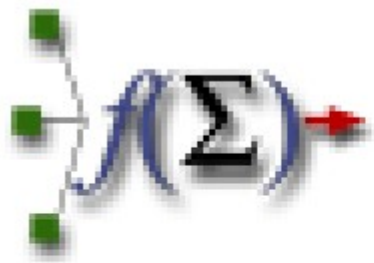
Nowe podejście - uwzględnianie wszystkich danych ze zdarzenia jednocześnie

Przykład - Icarus



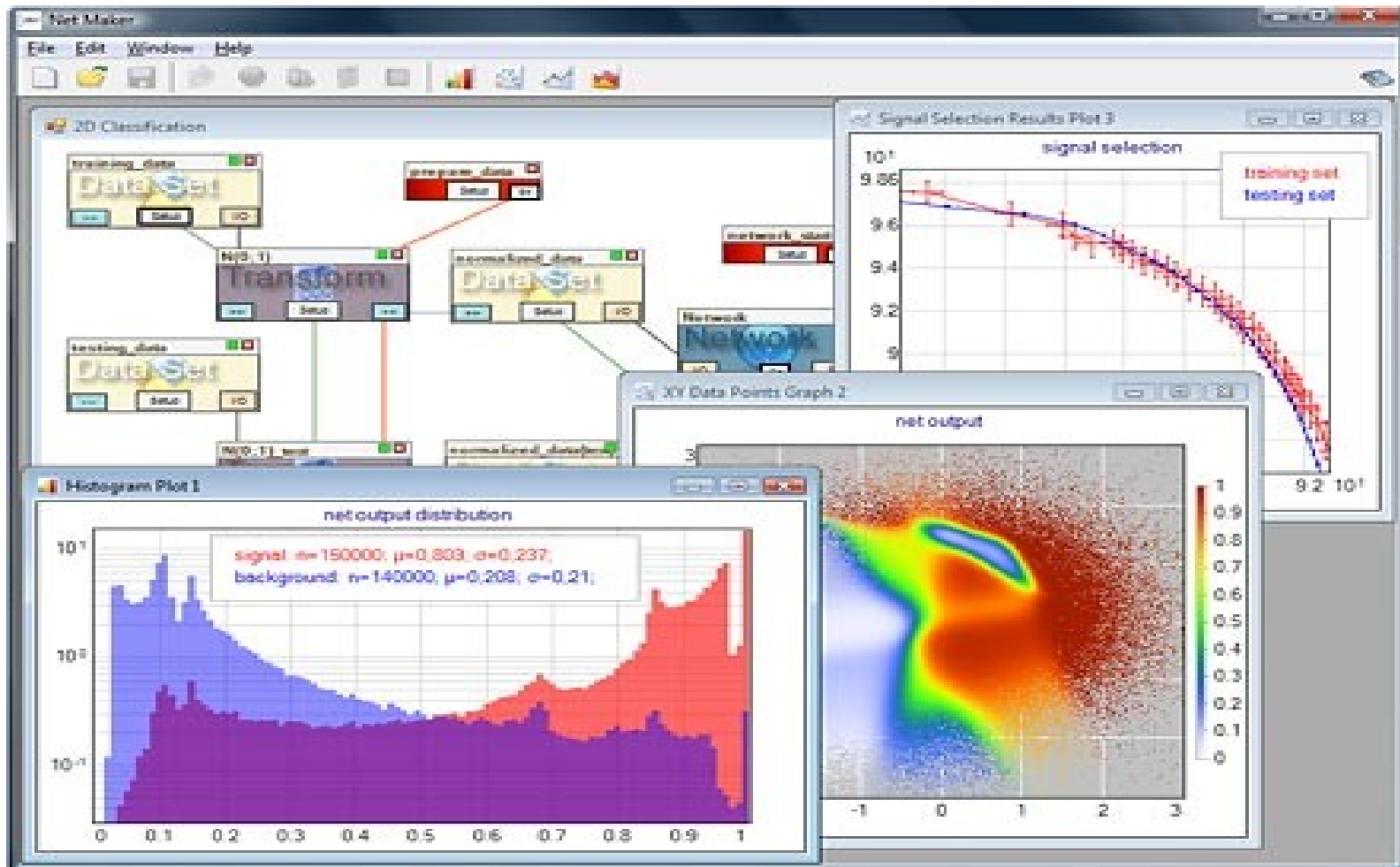
Wyniki - porównanie





NetMaker

neural networks simulator and designer



Dziękuję za uwagę!