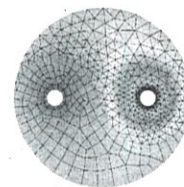




Wydawnictwo  
Naukowe  
AKAPIT



## IDENTYFIKACJA PLANU DOŚWIADCZENIA DLA APROKSYMACJI NEURONOWEJ

ROMA GÓRECKA  
JACEK PIETRASZEK  
ZBIGNIEW POLAŃSKI

### IDENTIFICATION OF THE EXPERIMENT PLAN FOR NEURON APPROXIMATION

Abstract

*The paper describes an application of the SDM method (sequence data modification) to searching for an optimum structure of training data set (experiment plan) in a statical neuron approximation.*

## 1. WSTĘP

### 1.1. Wprowadzenie

Od kilku lat w Samodzielnej Pracowni Zastosowań Informatyki Politechniki Krakowskiej z inicjatywy i pod kierunkiem Z. Polańskiego prowadzone są prace nad adaptacją sieci neuronowych do zagadnień związanych z planowaniem doświadczeń w technice. Istotnym problemem z tym związanym jest umiejętność oszacowania niepewności parametrów identyfikowanego statycznego obiektu badań, którego zachowanie jest aproksymowane siecią neuronową. Niniejsza praca jest fragmentem szerszych działań prowadzonych w ramach projektu badawczego KBN 8 T10C 003 15 "Integracja metod teorii eksperymentu i aproksymacji neuro-nowej w metrologii" i przedstawia wybrane wyniki uzyskane w trakcie badania aproksymatorów neuronowych. Z uwagi na ograniczoną objętość arty-

kułu zamieszczono tylko jeden przykład obliczeniowy, który przy swojej prostocie, nie zaciemniającej poglądości, ilustruje prezentowaną metodę.

### 1.2. Zarys metody SMD

Jednym z ważnych użytecznych zastosowań sieci neuronowych jest aproksymacja nieznanymi związków funkcyjnych dla wielow wejściowego  $(x_1, \dots, x_i)$  i jedno-wyjściowego  $(z)$  obiektu badań (Polański, 1984). Najczęściej dysponuje się dyskretnym zbiorem danych:

$$D_i = \left\{ \left\{ \left\{ x_1^{(1)}, \dots, x_i^{(1)} \right\}, z^{(1)} \right\}, \dots, \left\{ \left\{ x_1^{(n)}, \dots, x_i^{(n)} \right\}, z^{(n)} \right\} \right\} \quad (1)$$

gdzie:  $x_k^{(u)}$  – wartości argumentów funkcji ( $k = 1, 2, \dots, i$ ;  $u = 1, 2, \dots, n$ ),  $z^{(u)}$  – wartości nieznannej funkcji aproksymowanej  $f$  uzyskane w wyniku np. pomiarów,  $i$  – liczba wielkości wejściowych,  $n$  – liczba elementów

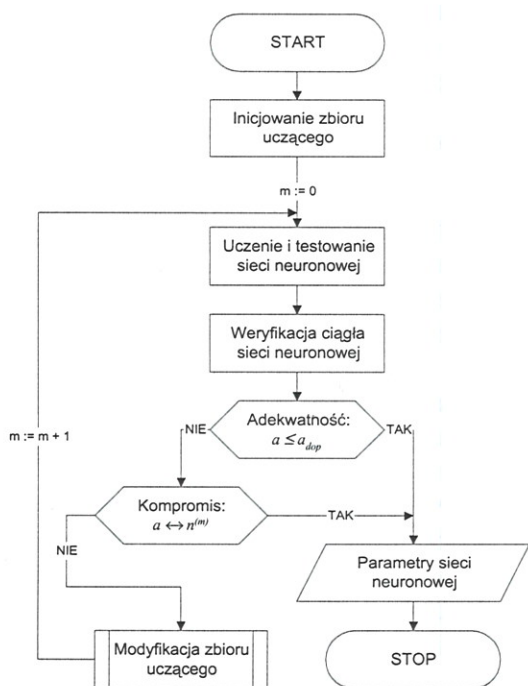
R. Górecka, Politechnika Krakowska, Wydział Mechaniczny, al. Jana Pawła II 37, 31-864 Kraków



zbioru danych będąca jednocześnie liczbą doświadczeń. Wiadomo, że problem identyfikacji (Mańczak i Nahorski, 1983) polega na wyborze odpowiedniej struktury modelu neuronowego, a następnie na estymacji parametrów modelu poprzez minimalizację funkcji celu bazującej na definicji błędu pomiędzy wyjściem obiektu a wyjściem modelu. Dalsze działania prowadzą do uzyskania modelu zgodnie z teorią aproksymacji (Kucharczyk, 1987). Elementem nietypowym w stosunku do zwykłej procedury aproksymacyjnej jest zastosowanie nieliniowego aproksymatora neuronowego charakteryzującego się dużą liczbą parametrów zwanych *wagami synaptycznymi* oraz wdrożenie specyficznych procedur identyfikacji parametrów modelu zwanych *uczeniem sieci* (Hertz i in., 1995, Korbicz i in., 1994).

Celem badań optymalizujących strukturę danych w powyższym zbiorze jest minimalizacja liczby danych stanowiących zbiory uczące (Górecka i in., 1999) i testujące sieć neuronową przy zachowanym podstawowym warunku: uzyskaniu założonej dokładności aproksymacji neuronowej. Obszarem zastosowań, w którym redukcja liczby danych ma sens są przypadki cechujące się znacząco dużymi kosztami pozyskiwania danych m.in. badania doświadczalne. Każda redukcja danych powoduje utratę informacji. Autorzy zaproponowali (Polański i in., 1999a, Polański i in., 1999b) procedurę sekwencyjną (iteracyjną) modyfikowania struktury danych prowadzącą do kompromisu akceptowanego arbitralnie przez badacza.

Koncepcja metody SMD (sekwencyjnej modyfikacji danych) ewoluowała w trakcie ostatnich dwóch lat.



Rysunek 1. Schemat algorytmu metody sekwencyjnej modyfikacji danych.

Początkowa propozycja została przedstawiona w roku 1997 (Polański i in., 1997), a później była rozwijana i prezentowana m.in. w materiałach seminarium *NeuroMet* w roku 1998 (Polański i Pietraszek, 1998), a jej skonkretyzowana postać w roku 1999 (Polański i in., 1999). Ogólny algorytm metody przedstawiony jest na rysunku 1.

## 2. CEL, PRZEDMIOT I ZAKRES BADAŃ

Typowe zastosowanie planowania doświadczeń stawia sobie za cel identyfikację funkcji obiektu badań przy jak najmniejszej liczbie układów (punktów) w przestrzeni wejść. Opisywane tutaj zagadnienie polega zaś na podejściu odwrotnym:

*jak rozmieścić zadaną liczbę układów (punktów) w przestrzeni wejść tak, aby identyfikowana funkcja obiektu badań minimalizowała przyjęte kryterium błędu aproksymacji.*

Zbiór powyższych układów w przestrzeni wejść obiektu badań definiuje zbiór uczący (zob. wzór (1)). Celem badań było określenie optymalnej struktury danych zbioru uczącego sieć neuronową, czyli właśnie rozmieszczenia tychże układów. Przedmiotem badań był pięciopunktowy zbiór danych w przestrzeni liniowej  $\mathbf{R}^2$  odwzorowujący kwadratową funkcję referencyjną, który stanowił zbiór uczący dla klasy perceptronów wielowarstwowych. Ujmując to tradycyjnym nazewnictwem teorii eksperymentu celem było wyznaczenie unormowanych wartości wielkości wejściowych dla planu doświadczenia o dwóch wielkościach wejściowych i pięciu układach.

Zakres pracy – zgodnie z algorytmem metody SMD – obejmował uczenie sieci neuronowych przy pomocy różnych wariantów zbioru uczącego, wyznaczanie błędów aproksymacji nad zbiorem weryfikującym (referencyjnym), statystyczną analizę uzyskanych wyników i wyznaczenie optymalnej struktury zbioru uczącego.

## 3. METODYKA I TECHNIKA BADAŃ

### 3.1. Metodyka badań

Jako funkcję referencyjną przyjęto funkcję kwadratową:

$$z = x_1^2 + x_2^2 \quad x_1, x_2 \in [-1, 1] \quad (2)$$

Zbiór weryfikujący został ustalony poprzez równy podział dziedziny funkcji w odstępnie 0,01 dla każdej wielkości wejściowej, co dało w efekcie 40 401 punktów weryfikujących:

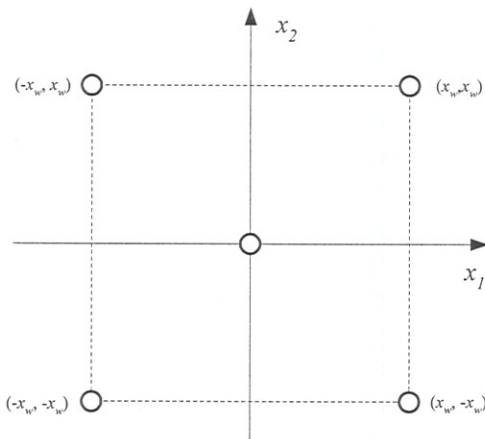


$$W = \left\{ \begin{array}{l} (x_{1/i}, x_{2/j}, f(x_{1/i}, x_{2/j})) \\ x_{1/i} = -1 + 0.01 \cdot i, \quad x_{2/j} = -1 + 0.01 \cdot j, \quad i, j = 0K 200 \end{array} \right\} \quad (3)$$

Zbiór powyższy nie jest odpowiednikiem zbioru uczącego D (por. (1)), lecz zbiorem bazowym dla obliczenia wartości globalnego kryterium błędu aproksymacji. Dla argumentów zbioru uczącego D (por. (1)) (planu doświadczenia) ustalono arbitralnie konfigurację bazową złożoną z pięciu punktów w przestrzeni wejść obiektu (zobacz tablica 1 i rysunek 2), sparametryzowaną przy pomocy wielkości  $x_w$ , której wartości mogą zmieniać się w przedziale lewostronnie otwartym  $(0, 1]$ .

Tablica 1. Identyfikowany plan doświadczenia.

| L.p. | $x_1$  | $x_2$  |
|------|--------|--------|
| 1.   | $-x_w$ | $-x_w$ |
| 2.   | $-x_w$ | $x_w$  |
| 3.   | $x_w$  | $-x_w$ |
| 4.   | $x_w$  | $x_w$  |
| 5.   | 0      | 0      |



Rysunek 2. Diagram rozmieszczenia układów planu.

Powszechnie do oceny jakości aproksymacji używane są następujące wskaźniki błędów:

- błąd maksymalny

$$MaxErr = \max_{i,j=0 \dots 200} |P_{w,n}(x_{1/i}, x_{2/j}) - f(x_{1/i}, x_{2/j})| \quad (4)$$

$$(x_{1/i}, x_{2/j}, f(x_{1/i}, x_{2/j})) \in W$$

- błąd przeciętny

$$AvgErr = \frac{1}{201 \cdot 201} \sum_{i=0}^{201} \sum_{j=0}^{201} |P_{w,n}(x_{1/i}, x_{2/j}) - f(x_{1/i}, x_{2/j})| \quad (5)$$

- błąd średniokwadratowy

$$RMS = \sqrt{\frac{1}{4} \int_{-1}^1 \int_{-1}^1 (P_{w,n}(x_1, x_2) - f(x_1, x_2))^2 dx_1 dx_2} \quad (6)$$

Jednocześnie jako kryterium jednorodności rozkładu lokalnego błędu aproksymacji przyjęto stosunek błędu maksymalnego  $MaxErr$  do błędu średniokwadratowego  $RMS$ .

We wzorach (4)-(6)  $f(x)$  jest funkcją referencyjną, natomiast  $P_{w,n}(x)$  jest funkcją opisującą odpowiedzi perceptronu o  $w$  warstwach ukrytych i  $n$  neuronach w każdej warstwie ukrytej. Przyjęto sigmoidalną funkcję aktywacji dla wszystkich neuronów. Uzyskane wyniki poddano analizie statystycznej (statystyki opisowe, analiza wariancji), a następnie przy pomocy regresji kwadratowej wyznaczono optymalne wartości wielkości wejściowej dla poszczególnych kryteriów błędu i sieci neuronowych.

Jako kryterium oceny aproksymacji przyjęto wskaźnik błędu maksymalnego  $MaxErr$ . Jednocześnie uznano, że prócz poszukiwania optymalnego rozmieszczenia wartości unormowanych planu powinno się także określić najlepszą topologię sieci. Z uwagi na złożoność zagadnienia w niniejszej pracy przyjęto, że perceptron posiada jedną warstwę ukrytą, natomiast optymalizowana jest liczba neuronów.

Poszukiwanie planu doświadczenia sprowadziło się wówczas do minimalizacji funkcji kryterium błędu maksymalnego nad dwuwymiarową przestrzenią, gdzie jedną współrzędną jest unormowana wartość planu  $x_w$ , drugą współrzędną – liczba neuronów sieci  $n$ .

Tak postawione zagadnienie samo w sobie jest typowym obiektem badań, którego zachowanie można opisać funkcją:

$$F : X_w \times N \rightarrow \mathcal{V} \quad (7)$$

gdzie:

- $X_w$  – podzbiór liczb rzeczywistych zawierających rozważane wartości parametru  $x_w$  poszukiwanego planu doświadczenia,
- $N$  – podzbiór liczb naturalnych zawierający rozważaną liczbę neuronów aproksymatora.

Przedstawienie funkcji  $F$  w postaci analitycznej nie jest możliwe. Badając ją sięgnięto więc po metody teorii eksperymentu.

Przeprowadzono dziewięć symulacji numerycznych zgodnie z trójwartościowym planem Boxa-Behnkena PS/DS: 32. (zobacz Polański, 1985, s. 204). Przyjęte dopuszczalne zakresy zmian wartości obejmowały:

- dla wartości parametru  $x_w$ : 0,2...0,6

- dla liczby neuronów  $n$ : 2...30

co w efekcie dało następującą dyskretyzację:  $x_w = 0,2$ ; 0,4; 0,6 oraz  $n = 2, 16, 30$ .

### 3.2. Technika badań

Symulatorem neuronowym stosowanym w trakcie badań był program *Statistica Neural Networks* (StatSoft, 1998). Przebiegi robocze programu realizowane były na typowym komputerze typu PC. Zbiory trenin-gowy i weryfikujący dostarczane były programowi w postaci formatowanych plików tekstowych zgodnych z wymaganiami programu. Wyniki wydawane były tak-że w formie plików tekstowych.

Uzyskane tekstowe pliki wynikowe były następnie przetwarzane przy pomocy niewielkich programów stworzonych w zespole badawczym oraz programu *Statistica* (StatSoft, 1997), a skonsolidowane wyniki (wskaźniki błędów) składowane w formatowanych plikach tekstowych.

Wskaźniki błędu zostały następnie opracowane sta-tystycznie przy pomocy programu *Statistica*. Wyni-kiem były rutynowe analizy statystyczne oraz analizy wariancji.

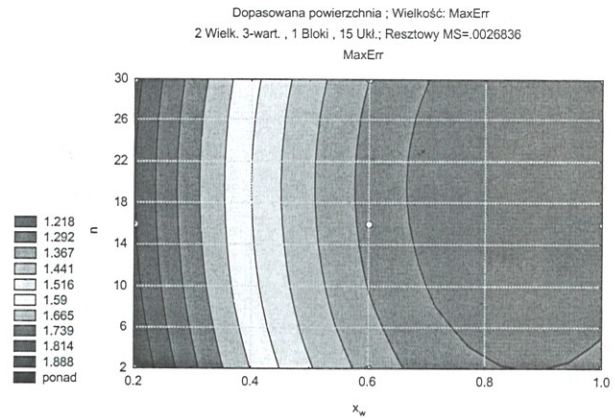
### 4. WYNIKI BADAŃ

Uzyskane wyniki przedstawiono poniżej w tablicy 2. W tablicy tej  $x_w$  to przyjęta wartość unormowana dla planu doświadczenia, natomiast  $n$  to liczba neuro-nów w warstwie ukrytej perceptronu.

Tablica 2. Wyniki symulacji.

| L.p. | $x_w$ | $n$ | MaxErr | AvgErr | RMS    | MaxErr / RMS |
|------|-------|-----|--------|--------|--------|--------------|
| 1    | 0.2   | 16  | 1.9220 | 0.6001 | 0.7328 | 2.62         |
| 2    | 0.2   | 2   | 1.9204 | 0.6037 | 0.7340 | 2.62         |
| 3    | 0.2   | 30  | 1.9276 | 0.6030 | 0.7338 | 2.63         |
| 4    | 0.6   | 16  | 1.2838 | 0.2951 | 0.3822 | 3.36         |
| 5    | 0.6   | 16  | 1.2817 | 0.3130 | 0.4016 | 3.19         |
| 6    | 0.6   | 16  | 1.2806 | 0.2210 | 0.3163 | 4.05         |
| 7    | 0.6   | 16  | 1.2852 | 0.3165 | 0.4060 | 3.17         |
| 8    | 0.6   | 16  | 1.2812 | 0.2810 | 0.3649 | 3.51         |
| 9    | 0.6   | 16  | 1.2805 | 0.2546 | 0.3374 | 3.79         |
| 10   | 0.6   | 16  | 1.2846 | 0.2862 | 0.3727 | 3.45         |
| 11   | 0.6   | 2   | 1.2840 | 0.3182 | 0.4080 | 3.15         |
| 12   | 0.6   | 30  | 1.2808 | 0.2123 | 0.3118 | 4.11         |
| 13   | 1     | 16  | 1.0714 | 0.4200 | 0.5272 | 2.03         |
| 14   | 1     | 2   | 1.3398 | 0.5483 | 0.6755 | 1.98         |
| 15   | 1     | 30  | 1.1905 | 0.4719 | 0.5881 | 2.02         |

Powyższe wyniki aproksymowano przy pomocy programu *Statistica* w celu wyznaczenia minimum wskaźnika błędu maksymalnego. Uciągloną po-wierzchnię aproksymującą przedstawiono na rysunku 3. Znalezione minimum ma wartość  $MaxErr = 1,1434$  i znajduje się w punkcie:  $x_w = 0,87861$  i  $n = 19,75$ . Oczywiście ostateczna liczba neuronów musi być za-okrąglona albo do wartości 19, albo do wartości 20.



Rysunek 3. Wskaźnik błędu maksymalnego.

Tablica 3. Wyniki symulacji podczas weryfikacji.

| L.p. | $x_w$ | $n$ | MaxErr | AvgErr | RMS    | MaxErr/RMS |
|------|-------|-----|--------|--------|--------|------------|
| 1    | 0.8   | 19  | 0.7240 | 0.2433 | 0.3017 | 2.40       |
| 2    | 0.8   | 20  | 0.7250 | 0.2371 | 0.3027 | 2.39       |
| 3    | 0.8   | 21  | 0.9759 | 0.3254 | 0.4018 | 2.43       |
| 4    | 0.85  | 19  | 0.7754 | 0.2824 | 0.3478 | 2.23       |
| 5    | 0.85  | 20  | 0.6353 | 0.2350 | 0.2955 | 2.15       |
| 6    | 0.85  | 21  | 0.6262 | 0.2441 | 0.2996 | 2.09       |
| 7    | 0.9   | 19  | 0.9530 | 0.3852 | 0.4758 | 2.00       |
| 8    | 0.9   | 20  | 1.0136 | 0.4072 | 0.5039 | 2.01       |
| 9    | 0.9   | 21  | 0.8895 | 0.3511 | 0.4359 | 2.04       |
| 10   | 0.95  | 19  | 1.9997 | 0.5540 | 0.6933 | 2.88       |
| 11   | 0.95  | 20  | 1.9997 | 0.5557 | 0.6947 | 2.88       |
| 12   | 0.95  | 21  | 1.9998 | 0.5521 | 0.6902 | 2.90       |

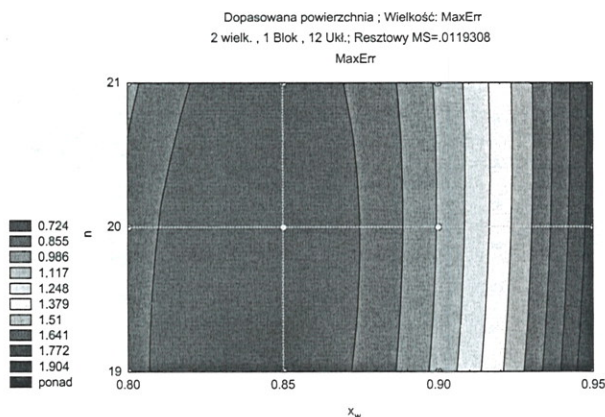
Uzyskane wyniki wymagają sprawdzenia, gdyż ich wartości zostały określone na podstawie aproksyma-cji. Przeprowadzono weryfikację poprzez ponowne zastosowanie opisanej procedury do zawężonego za-kresu wartości:  $x_w = 0,85...0,95$  i  $n = 19...21$ . Nieste-ty okazało się, że minimum znajduje poza tym obsza-rem i konieczne było poszerzenie go do kompletnej kombinacji wartości:  $x_w = \{0,80 ; 0,85 ; 0,90 ; 0,95\}$  i  $n = \{19, 20, 21\}$ . Uzyskane wyniki przedstawiono w tablicy 3.

Uzyskane wyniki wymagają sprawdzenia, gdyż ich wartości zostały określone na podstawie aproksyma-cji. Przeprowadzono weryfikację poprzez ponowne zastosowanie opisanej procedury do zawężonego za-kresu wartości:  $x_w = 0,85...0,95$  i  $n = 19...21$ . Nieste-ty okazało się, że minimum znajduje poza tym obsza-rem i konieczne było poszerzenie go do kompletnej kombinacji wartości:  $x_w = \{0,80 ; 0,85 ; 0,90 ; 0,95\}$  i  $n = \{19, 20, 21\}$ . Uzyskane wyniki przedstawiono w tablicy 3.

Ostatecznie jako rozwiązanie przyjęto:

- wartość unormowaną dla planu doświadczenia:  $x_w = 0,8407$ ,





Rysunek 4. Wskaźnik błędu maksymalnego w trakcie weryfikacji.

- liczba neuronów najlepiej aproksymującej sieci:  $n = 19 \vee 20$ .

Powyższe wartości minimalizują przyjęte kryterium błędu aproksymacji w założonym obszarze zmienności parametrów sieci.

## 5. WNIOSKI

Z zamieszczonych powyżej wyników można wysnuć następujące wnioski:

1. procedura postępowania jest skuteczna i efektywna – prowadzi do uzyskania rozwiązania,
2. założenie dotyczące kształtu funkcji wzorcowej wymaga weryfikacji po uzyskaniu wyników rzeczywistych pomiarów (nota bene: dotyczy to także tradycyjnej teorii eksperymentu),
3. procedura SMD umożliwia modyfikację planu doświadczenia w sposób uwzględniający posiadane już wyniki pomiarów – jest to swoista procedura poprawiania planu doświadczenia.

## LITERATURA

- Górecka R., Pietraszek J., Polański Z., 1999, Zastosowanie wielomianów Czebyszewa w aproksymacji neuronowej. *II Ogólnopolska Konferencja Naukowa Quality'99*, Kraków t. 2, 337-342.
- Hertz J., Krogh A., Palmer R.G., 1995, *Wstęp do teorii obliczeń neuronowych*, WNT, Warszawa.
- Korbicz J., Obuchowicz A., Uciński D., 1994, *Sztuczne sieci neuronowe, podstawy i zastosowania*, AOW PLJ, Warszawa.
- Mańczak K., Nahorski Z., 1983, *Komputerowa identyfikacja obiektów dynamicznych*, PWN, Warszawa.
- Polański Z., 1984, *Planowanie doświadczeń w technice*, PWN, Warszawa.
- Polański Z., Pietraszek J., 1998, Aproksymacja neuronowa w zastosowaniach ograniczających liczebność danych uczących, *Materiały II Seminarium NeuroMet'98 Zastosowanie sztucznych sieci neuronowych w symulacji i sterowaniu procesami metalurgicznymi*, ed., Kusiak J., Kraków, 58-66.
- Polański Z., Pietraszek J., Górecka R., 1999a, Aproksymacja neuronowa metodą sekwencyjnej modyfikacji danych, *Materiały III Seminarium NeuroMet'99 Zastosowanie sztucznych sieci neuronowych w symulacji i sterowaniu procesami metalurgicznymi*, ed., Kusiak J., Kraków, 21-37.
- Polański Z., Pietraszek J., Górecka R., 1999b, Optymalizacja struktury danych do aproksymacji neuronowej, *Materiały IV Szkoły-Konferencji Metrologia Wspomagana Komputerowo MWK'99*, Rynia k. Warszawy, t. 3, 11-16.
- Polański Z., Pietraszek J., Górecka R., Osocha P., 1997, Badania wpływu struktury danych na proces uczenia sieci neuronowej, *Materiały IV Szkoły-Konferencji Metrologia Wspomagana Komputerowo MWK'97*, Zegrze k. Warszawy, t. 2, 19--24.
- Kucharczyk P., (ed.), 1987, *Poradnik inżyniera. Matematyka*, WNT, Warszawa, t. 2, 429-457.
- StatSoft, Inc., 1998, *STATISTICA Neural Networks* [Computer program manual]. Tulsa, OK: StatSoft, Inc., 2300 East 14th Street, Tulsa, OK 74104, phone: (918) 749-1119, fax: (918) 749-2217, email: info@statsoftinc.com, WEB: http://www.statsoft.com
- StatSoft, Inc., 1997, *STATISTICA for Windows* [Computer program manual]. Tulsa, OK: StatSoft, Inc., 2300 East 14th Street, Tulsa, OK 74104, phone: (918) 749-1119, fax: (918) 749-2217, email: info@statsoftinc.com, WEB: http://www.statsoft.com