

ANDRZEJ PIEGAT

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

BARBARA WĄSIKOWSKA

Uniwersytet Szczeciński

MARCIN KORZEŃ

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

ZASTOSOWANIE SAMOUCZĄCEGO SIĘ TRZYPUNKTOWEGO MINIMODELU DO MODELOWANIA STOPY BEZROBOCIA W POLSCE

Wprowadzenie

W gospodarce każdego kraju istnieje wiele funkcyjnych, wielowymiarowych zależności ekonomicznych typu $y = f(x_1, x_2, \dots, x_n)$. W ostatnim dwudziestolecu coraz częściej dokonuje się identyfikacji zależności ekonomicznych z użyciem sztucznych sieci neuronowych¹. Neuronowe modele zależności funkcyjnych są modelami samouczącymi się, które potrafią dopasować powierzchnię neuronowego modelu do posiadanych próbek uczących – realizują więc proces

¹ J. Morajda, *Neural networks as predictive models in financial futures trading*, Fifth Conference on Neural Networks and Soft Computing, Zakopane 2000; A. Pokropińska, R. Scherer, *Financial prediction with neuro-fuzzy systems*, International Conference AISC 2008, Springer 2008; I. Rejer, *Metody modelowania wielowymiarowego systemu z użyciem metod sztucznej inteligencji na przykładzie bezrobocia w Polsce*, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin 2003; B. Wąsikowska, *Niestandardowe metody identyfikacji czynników determinujących produkcję sprzedaną przemysłu*, Uniwersytet Szczeciński, Szczecin 2007 (rozprawa doktorska); D. Witkowska, *Application of econometric models and neural networks to consumer price index prediction*, Fifth Conference on Neural Networks and Soft Computing, Zakopane 2000.

samoadaptacji. I tak przykładowo, opracowano model neuronowy² dla stopy bezrobocia w Polsce oraz neuronowy model³ dla produkcji sprzedanej przemysłu.

Cechą większości modeli ekonometrycznych, neuronowych i innych jest ich globalny charakter, to znaczy definiują one wartości zmiennej wyjaśnianej y w całej przestrzeni rozważań, zwykle prostokątnej, zawartej między maksymalnymi a minimalnymi wartościami poszczególnych zmiennych x_i . Globalność modelu to na pewno korzystna jego cecha. W praktyce jednak pociąga ona za sobą wiele negatywów, strojenie modelu globalnego jest bowiem znacznie trudniejsze niż „małego” modelu lokalnego. Dodatkowo średni błąd modelu globalnego może być akceptowalnie mały, jednak lokalnie model ten w niektórych podzakresach przestrzeni rozważań może wykazywać znacznie wyższe błędy (w innych znacznie niższe) niż średnio.

Dodatkowo, w miarę napływu nowych danych, globalny model – nastrojony wcześniej posiadany zestawem danych – dezaktualizuje się i wymaga ciągłego dostrajania, co może być kłopotliwe, jeśli jego użytkownicy, na przykład pracownicy firmy, nie potrafią tego zrobić. Stąd, ze względów praktycznych, w wielu wypadkach korzystne może być stosowanie szybkich, samouczących się modeli lokalnych. Modele te dostarczają odpowiedzi na interesujące nas zapytanie na podstawie lokalnego modelu. Model ten dokonuje ekstrakcji wiedzy potrzebnej do odpowiedzi na zapytanie na podstawie próbek (danych) okrążających zapytanie. Jeżeli rozpatrywana jest pewna interesująca nas trójwymiarowa zależność $y = f(x_1, x_2)$, to zapytanie i odpowiedź mają formę:

$$\begin{aligned} f(x_1^*, x_2^*) &= ? - \text{zapytanie,} \\ y^* &= f(x_1^*, x_2^*) - \text{odpowiedź na zapytanie} \end{aligned} \quad (1)$$

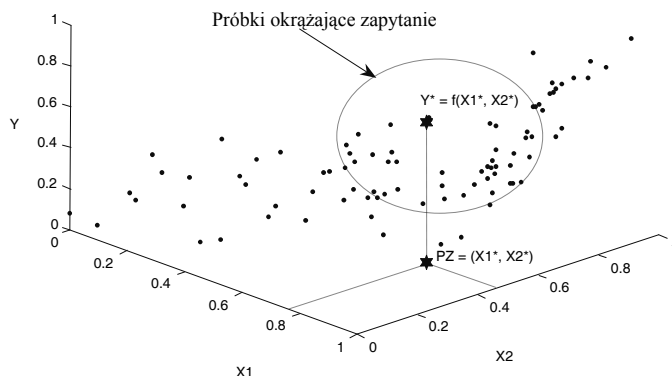
gdzie:

- y^* – wartość liczbowa,
- x_1^*, x_2^* – liczbowe wartości zmiennych X_1, X_2 ,
- $y = f(x_1, x_2)$ – zależność,
- PZ: $\{x_1^*, x_2^*\}$ – punkt zapytania.

Rysunek 1 ilustruje pojęcia związane z zapytaniem i odpowiedzią, jakiej oczekuje się od modelu lokalnego.

² I. Rejer, *Metody modelowania...*

³ B. Wąsikowska, *Niestandardowe metody identyfikacji...*



Rys. 1. Ilustracja pojęć związanych z zapytaniem do modelu

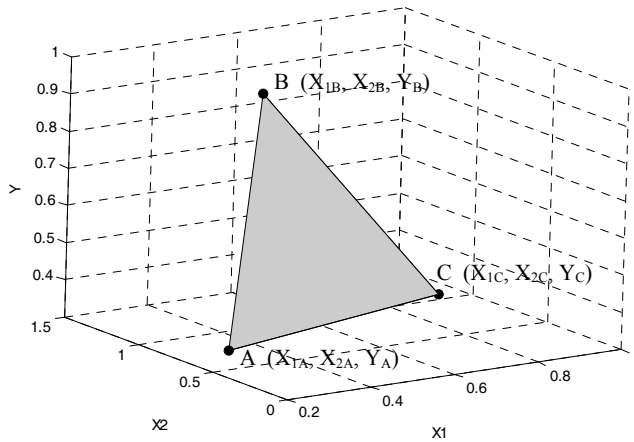
Źródło: opracowanie własne.

Aby znaleźć odpowiedź na zapytanie, należy stworzyć mały, lokalny model o skończonym zakresie, obejmujący swym zasięgiem jedynie pobliskie próbki okrażające zapytanie. W dalszej części artykułu model taki nazywany będzie minimodelem (MM). Minimodele mogą mieć różne kształty geometryczne. W dalszym ciągu przedstawiona zostanie koncepcja minimodelu trójkątnego w przestrzeni 3D, któremu w przestrzeni nD, o wyższym wymiarze, odpowiadają różne ostrosłupy. W przypadku przestrzeni 4D minimodel jest czworościanem o trójkątnych ścianach, zwykle nierównobocznych.

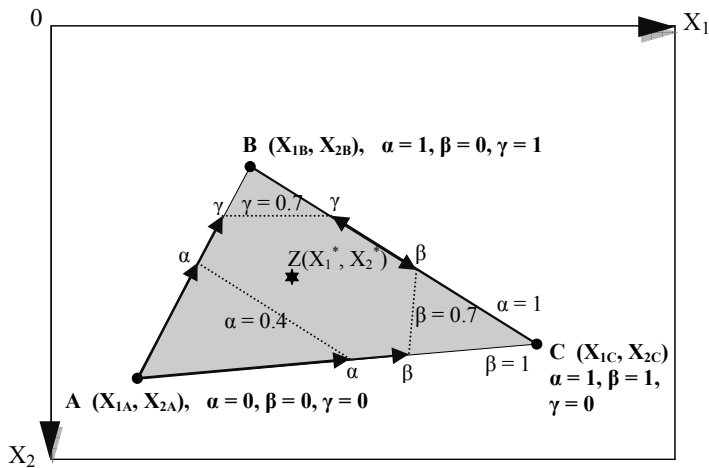
1. Trójkątny, liniowy, ograniczony minimodel w przestrzeni 3D

Koncepcja minimodelu – oraz odpowiednie wzory – opracowana została przez A. Piegata. Oparta jest ona na kontekstowym, ograniczonym układzie współrzędnych α , β . Układ ten jest oparty na trzech punktach narożnych A, B, C (rysunki 2a i 2b).

a)



b)



Rys. 2. Trójkątny układ kontekstowy $\{\alpha, \beta, \gamma\}$ – ilustracja wprowadzonych oznaczeń $\alpha \in [0, 1]$, $\beta \in [0, 1]$, $\gamma \in [0, 1]$, Z – przykładowy punkt zapytania

Źródło: opracowanie własne.

W kontekście minimodelu $\{A, B, C\}$ rozpatrywane są tylko punkty $\{X_1, X_2\}$ leżące w jego wnętrzu. Czy dany punkt $\{X_1, X_2\}$ leży wewnątrz minimodelu łatwo sprawdzić, każdy taki punkt musi bowiem spełniać jednocześnie trzy warunki: $\alpha \in [0, 1]$, $\beta \in [0, 1]$, $\gamma \in [0, 1]$. Wartości współrzędnych kontekstowych $\alpha, \beta,$

γ można obliczyć ze wzorów (2) uzyskanych na zależnościach geometrycznych dla dowolnego punktu $\{X_1, X_2\}$:

$$\begin{aligned}\alpha &= (x_1, x_2) = \frac{(x_1 - x_{1A})(x_{2C} - x_{2B}) - (x_2 - x_{2A})(x_{1C} - x_{1B})}{(x_{1B} - x_{1A})(x_{2C} - x_{2B}) - (x_{2B} - x_{2A})(x_{1C} - x_{1B})}, \\ \beta &= (x_1, x_2) = \frac{(x_1 - x_{1A})(x_{2B} - x_{2A}) - (x_2 - x_{2A})(x_{1B} - x_{1A})}{(x_{1C} - x_{1A})(x_{2B} - x_{2A}) - (x_{2C} - x_{2A})(x_{1B} - x_{1A})}, \\ \gamma &= (x_1, x_2) = \frac{(x_1 - x_{1A})(x_{2C} - x_{2A}) - (x_2 - x_{2A})(x_{1C} - x_{1A})}{(x_{1B} - x_{1A})(x_{2C} - x_{2A}) - (x_{2B} - x_{2A})(x_{1C} - x_{1A})}\end{aligned}\quad (2)$$

Znając współrzędne wierzchołków minimodelu (MM), łatwo obliczyć wartość zmiennej objaśnianej y dla punktu zapytania o współrzędnych kontekstowych α, β :

$$y = y_A + (y_B - y_A) \alpha + (y_C - y_B) \beta \quad (3)$$

2. Skrótowy opis algorytmu uczenia minimodelu

Celem działania minimodelu jest obliczenie odpowiedzi $y^* = f(x_1^*, x_2^*)$ dla zadanego, interesującego nas punktu zapytania PZ: $\{x_1^*, x_2^*\}$. Uczeniu podlegają wszystkie współrzędne punktów narożnych A, B, C minimodelu.

Kroki algorytmu uczenia minimodelu

Krok 1

Zainicjowanie startowych wartości wierzchołków A, B, C: (X_{1A}, X_{2A}, Y_A) , (X_{1B}, X_{2B}, Y_B) , (X_{1C}, X_{2C}, Y_C) tak, aby punkt zapytania PZ: $\{x_1^*, x_2^*\}$ znajdował się wewnątrz MM (warunek: $\alpha \in [0, 1]$, $\beta \in [0, 1]$, $\gamma \in [0, 1]$). Sprawdzenie, czy liczba próbek uczących n wewnątrz minimodelu przewyższa minimalną liczbę n_{min} zależną od wymiarowości MM (tutaj 3) oraz błędu pomiarowego próbek. Przy bardzo małym błędzie próbek minimalna liczba próbek $n_{min} = 3$ lub 4, przy większych błędach pomiaru n_{min} należy zwiększyć. Jeżeli któryś z warunków wymaganych w kroku 1 nie jest spełniony, należy wygenerować nowe wartości startowe współrzędnych punktów A, B, C. Dla wszystkich m próbek znaj-

dujących się wewnątrz MM obliczany jest średni błąd bezwzględny B_s według wzoru (4):

$$B_s = \frac{1}{n} \sum_{i=1}^n |y_{pi} - y_i| \quad (4)$$

gdzie:

- B_s – średni błąd bezwzględny modelu,
- n – liczba próbek w zbiorze uczącym,
- y_{pi} – współrzędna y próbki i ,
- y_i – wartość y obliczona dla próbki i według MM.

Jeśli $B_s \geq B_{s\min}$, to następuje zakończenie procesu uczenia. Jeżeli nie, należy przejść do kroku drugiego.

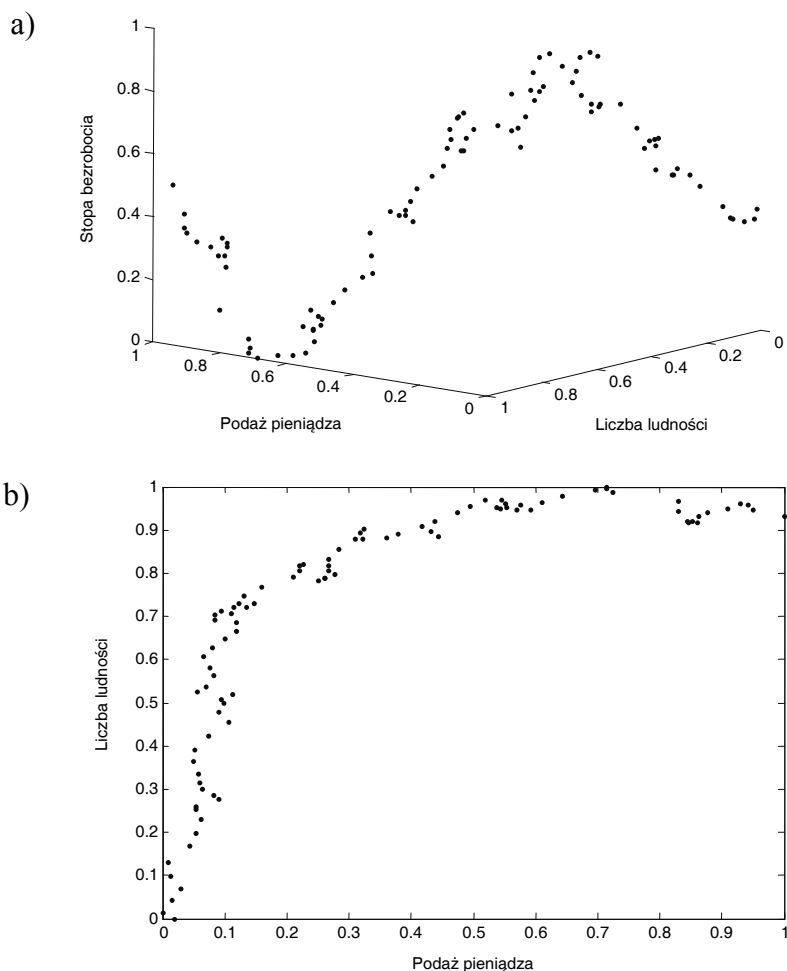
Krok 2

Zmiana aktualnej pozycji punktów A, B, C minimodelu tak, aby zmniejszyła się liczba próbek m znajdujących się wewnątrz MM oraz zmniejszył się średni błąd bezwzględny B_s dla próbek objętych minimodelem przy jednoczesnym spełnieniu warunku minimalnej liczby próbek $m \geq n_{\min}$. Następnie następuje sprawdzenie minimalnego błędu MM. Jeżeli $B_s < B_{s\min}$, to należy powtórzyć krok 2, a jeżeli $B_s \geq B_{s\min}$, to następuje zakończenie obliczeń.

3. Przykład użycia minimodelu do problemu bezrobocia w Polsce

Problem bezrobocia w Polsce badany był między innymi w pracy I. Rejer⁴. Badania wykazały, że najbardziej istotnymi czynnikami wyjaśniającymi wartość *stopy bezrobocia* w danym miesiącu w Polsce (y) w latach 1991–1999 (okres 96 miesięcy) były: X_1 – *podaż pieniądza* (w danym miesiącu) oraz X_2 – *liczba ludności* w Polsce (w danym miesiącu). Na rysunku 3 przedstawiony został rozkład wszystkich próbek uczących w przestrzeni $X_1 * X_2 * Y$ (rysunek 3a) oraz $X_1 * X_2$ (rysunek 3b).

⁴ I. Rejer, *Metody modelowania...*

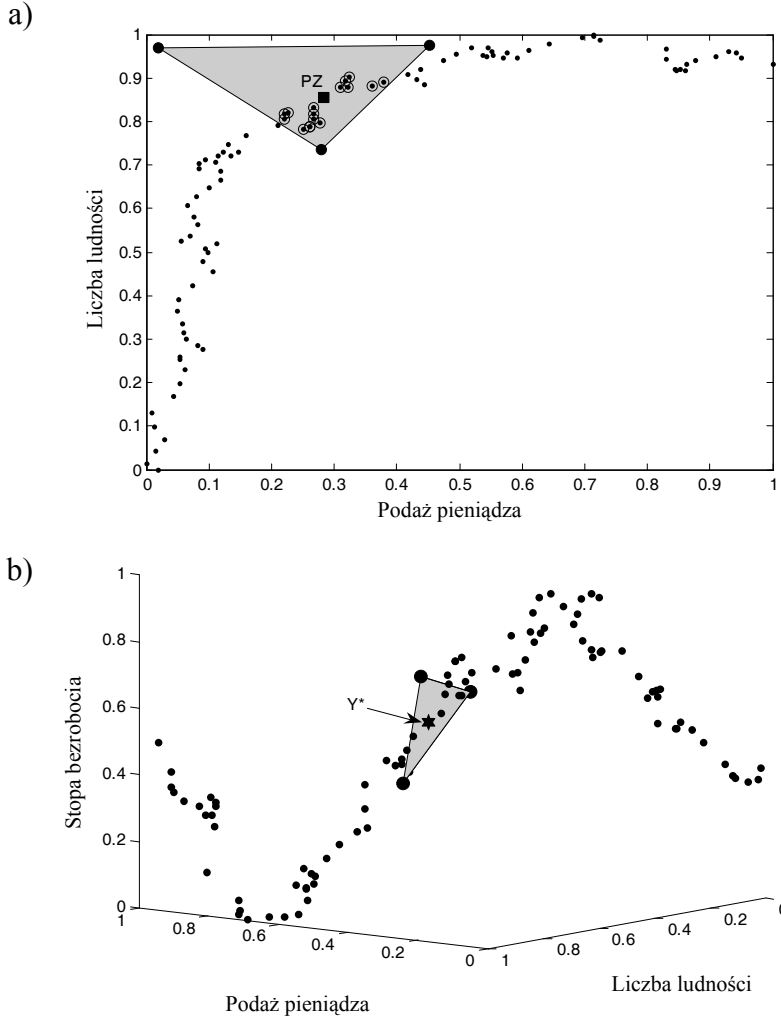


Rys. 3. Rozkład rzeczywistych próbek w przestrzeni $X_1 * X_2 * Y$ (rysunek 3a) oraz $X_1 * X_2$ (rysunek 3b)

Źródło: opracowanie własne.

Jak widać na rysunku 3, próbki w małym stopniu pokrywają przestrzeń wejść X_1 i X_2 . Stąd żaden model nie jest w stanie dać wiarygodną odpowiedź na dowolne zapytanie $\{x_1^*, x_2^*\}$. Wiarygodną odpowiedź można uzyskać tylko wówczas, gdy punkt zapytania (PZ) leży w pobliżu próbek uczących. Spełnienie tego warunku umożliwiają właśnie minimodele, natomiast często nie spełniają go modele globalne. Na rysunku 4 pokazano położenie przykładowego punktu

zapytania $\{x_1^*, x_2^*\} = \{0,28, 0,86\}$, startowy (rysunek 4a) i końcowy (rysunek 4b) minimodel uzyskany w wyniku procesu uczenia.



Rys. 4. Punkt zapytania PZ: $\{0,28, 0,86\}$, startowa (rysunek 4a) forma minimodelu i końcowa (rysunek 4b) forma minimodelu uzyskana w procesie uczenia

Źródło: opracowanie własne.

Na podstawie posiadanych próbek (dane statystyczne) minimodel ustalił odpowiedź (y^*), biorąc pod uwagę zapytanie (PZ):

$$\text{jeśli PZ: } \{x_1^*, x_2^*\} = \{0,28, 0,86\}, \text{ to } \{y^*\} = \{0,621\}.$$

Powstaje oczywiście kwestia dokładności odpowiedzi obliczeniowych udzielanych przez minimodel.

4. Testowanie dokładności wyników obliczeń minimodelu

Testowanie dokładności obliczeń realizowanych przez minimodel (MM) może być dokonane znaną metodą krzyżowej walidacji *leave one out*⁵, czyli „komplet minus jedna próbka”. W tym celu po kolei ze zbioru danych uczących zabierana jest jedna próbka o współrzędnych $\{x_{1i}, x_{2i}, y_i\}$, po czym znajdowana jest z użyciem minimodelu odpowiedź na zapytanie: „Ile wynosi wartość zmiennej y_{iMM} obliczona przez minimodel dla punktu zapytania o współrzędnych próbkki $\{x_{1i}, x_{2i}\}$?” Po znalezieniu predykcji y_{iMM} obliczany jest błąd wyniku dostarczonego przez minimodel według wzoru 5:

$$|e_i| = |y_i - y_{iMM}| \quad (5)$$

gdzie:

- $|e_i|$ – błąd bezwzględny modelu,
- y_i – rzeczywiste wartości wyjściowe ($i = 1, \dots, n$),
- y_{iMM} – wyznaczone przez minimodel wartości wyjściowe ($i = 1, \dots, n$).

Następnie obliczany jest średni błąd bezwzględny według wzoru 6:

$$B_s = \frac{1}{n} \sum_{i=1}^n |y_i - y_{iMM}| \quad (6)$$

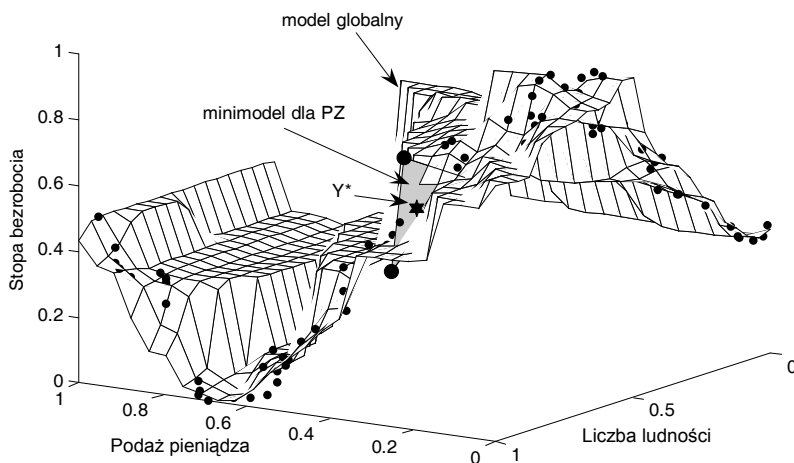
gdzie:

- B_s – średni błąd bezwzględny modelu,
- n – liczba próbek w zbiorze uczącym,
- y_i – rzeczywiste wartości wyjściowe ($i = 1, \dots, n$),
- y_{iMM} – wyznaczone przez minimodel wartości wyjściowe ($i = 1, \dots, n$).

⁵ S. Geissler, *Predictive inference*, Chapman and Hall, New York 1993.

Testowanie dokładności obliczeń minimodelu dla problemu bezrobocia wykazało, że średni błąd bezwzględny obliczeń dla wszystkich posiadanych 96 próbek wyniósł 0,02729. Dla porównania, bardzo dokładny model neuronowy o charakterze globalnym miał średnią wartość błędu równą 0,037501.

Na rysunku 5 przedstawiono powierzchnię globalnego modelu neuronowego oraz – dla porównania – przykładowy minimodel uzyskany w wyniku procesu uczenia dla punktu zapytania PZ: $\{0,28, 0,86\}$.



Rys. 5. Globalny model zależności $y = f(x_1, x_2)$ uzyskany za pomocą sztucznej sieci neuronowej i przykładowy minimodel uzyskany dla zapytania PZ: $\{0,28, 0,86\}$

Źródło: opracowanie własne.

Wnioski

W artykule przedstawiono koncepcję ograniczonego, lokalnego, trójkątnego minimodelu, którego celem jest udzielenie – na podstawie próbek zależności $y = f(x_1, x_2)$, jaką dysponujemy w bazie wiedzy – odpowiedzi na zapytania obliczeniowe. Badania wykazały, że dokładność obliczeń realizowanych przez minimodel jest wysoka, a czas obliczeń krótki. Uzyskana wysoka dokładność jest wynikiem lokalności minimodelu. Dalsze prace autorów idą w kierunku opracowania nieliniowych minimodeli o jeszcze wyższej dokładności obliczeń.

Literatura

- Geissler S., *Predictive inference*, Chapman and Hall, New York 1993.
- Morajda J., *Neural networks as predictive models in financial futures trading*, Fith Conference on Neural Networks and Soft Computing, Zakopane 2000.
- Pokropińska A., Scherer R., *Financial prediction with neuro-fuzzy systems*, International Conference AISC 2008, Springer 2008.
- Rejer I., *Metody modelowania wielkowymiarowego system z użyciem metod sztucznej inteligencji na przykładzie bezrobocia w Polsce*, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin 2003.
- Wąsikowska B., *Niestandardowe metody identyfikacji czynników determinujących produkcję sprzedaną przemysłu*, rozprawa doktorska, Wydział Nauk Ekonomicznych i Zarządzania, Uniwersytet Szczeciński, Szczecin 2007.
- Witkowska D., *Application of econometric models an neural networks to consumer price index prediction*, Fith Conference on Neural Networks and Soft Computing, Zakopane 2000.

**ANWENDUNG DES SELBSTLERNENDEN, 3-PUNKT- MINIMODELS
ZUR MODELLIERUNG DER ARBEITSLOSENQUOTE IN POLEN****Zusammenfassung**

Gegenstand des Artikels ist die Konstruktionsmethode des lokalen Minimodels das eine beträchtliche Lerngeschwindigkeit besitzt. Das Model dient hier zur Berechnung schneller Antworten an Anfragen betreffens der Arbeitslosenquote aufgrund der Daten. Die Anzahl der Daten muss nicht konstant sein und kann mit der Zeit vergrößert werden, falls neue Daten gewonnen werden. Ein großer Vorteil des Minimodels ist seine Fähigkeit zur Interpolation in den Informationslücken in den Daten. Diese Fähigkeit unterscheidet das vorgeschlagene Minimodel von der KNN-Methode (Methode der nächsten Nachbarn). Ausser dem Erklären der Minimodelsidee wurde in dem Artikel ein Beispiel der Anwendung zur Modellierung der Arbeitslosenquote aufgrund statistischer Daten gezeigt. Die Qualität des Minimodels wurde mit der bekannten „leave one out“ Methode geprüft.

Übersetzt von Andrzej Piegat

