

Monika Hadaś

Akademia Ekonomiczna w Katowicach

ZASTOSOWANIE SIECI FALKOWO-NEURONOWEJ DO PREDYKCJI EKONOMICZNYCH SZEREGÓW CZASOWYCH

Tematyka związana ze sztucznymi sieciami neuronowymi należy do interdyscyplinarnej dziedziny badań powiązanej m.in. z biocybernetyką, elektroniką, matematyką stosowaną, automatyką, a nawet medycyną. Sztuczne sieci neuronowe powstały na gruncie wiedzy o działaniu systemu nerwowego istot żywych i stanowią próbę wykorzystania zjawisk zachodzących w systemach nerwowych przy poszukiwaniu nowych rozwiązań technologicznych.

Sieci neuronowe należą do grupy układów charakteryzujących się sztuczną inteligencją. Jedną z podstawowych cech tej grupy jest zdolność uogólniania wiedzy, tzn. sieć nauczona na jednym zbiorze danych generuje właściwe wyniki przy podawaniu na jej wejście zbioru danych należących wprawdzie do tej samej grupy, ale nieuczestniczących w procesie uczenia.

W niniejszej pracy wykorzystujemy własności sztucznej sieci neuronowej i falek do przeprowadzenia predykcji szeregu czasowego. Prezentowana predykcja jest oparta na sieci falkowo-neuronowej, która jest odmianą sieci neuronowej, a w której tradycyjne funkcje aktywacji neuronów zastąpiono funkcjami falkowymi.

Ponieważ niemal wszystkie ekonomiczne szeregi czasowe są obarczone błędami, które wynikają z niedoskonałości przeprowadzonych obserwacji realizowanych na zmiennych czasowych, stosujemy dwie metody predykcji. Pierwszą metodą przeprowadzamy predykcję bezpośrednio na oryginalnym ekonomicznym szeregu czasowym bez robienia redukcji zakłóceń, a drugą metodą wykonujemy predykcję po wcześniejszym usunięciu zakłóceń, np. średnią ruchomą.

1. Podstawy teoretyczne falki i falkowej sieci neuronowej

1.1. Teoria falek

Falka jest funkcją, która dzięki operacji przesunięcia i przeskalowania generuje bazę ortonormalną przestrzeni $L^2(R)$. Funkcje te nazwano falkami, ponieważ są znakoziemne, czyli „faluja” i w sposób istotny różnią się od zera tylko na małym odcinku.

W przeciwieństwie do analizy fourierowskiej, analiza falkowa nie wyraża badanych funkcji poprzez wielomiany, ale poprzez specjalne funkcje – falki, które są generowane przez funkcję zwaną falką macierzystą $\psi(t)$:

$$\psi_{a,b}(t) = a^{-\frac{1}{2}} \psi\left(\frac{t-b}{a}\right) \quad a, b \in R \quad a > 0. \quad (1)$$

Uzyskane w ten sposób falki mają szereg interesujących własności. Można odnosić je zarówno do czasu, jak i do częstotliwości, dopuszczając bliższe związki między badaną funkcją a jej współczynnikami.

Transformatę falkową sygnału $f(t)$ definiuje się jako:

$$W_f(a, b) = c_\psi^{-\frac{1}{2}} \int f(t) \psi_{a,b}^*(t) dt, \quad (2)$$

gdzie $*$ jest operatorem sprzężenia, $c_\psi = 2\pi \int_0^\infty \frac{|\hat{\psi}(\omega)|}{\omega} d\omega$, gdzie $\hat{\psi}$ oznacza transformatę

Fouriera: $\hat{\psi}(\omega) = (2\pi)^{-\frac{1}{2}} \int e^{i\omega t} \psi(t) dt$.

Transformata odwrotna dla równania (2) wyrażona jest wzorem:

$$f(t) = c_\psi^{-\frac{1}{2}} \int \int W_f(a, b) \psi_{a,b}(t) \frac{da db}{a^2}. \quad (3)$$

W praktycznych zastosowaniach parametry a i b są wartościami dyskretnymi. Prowadzi to do definicji falki dyskretnej wyrażonej wzorem:

$$\psi_{j,k}(t) = 2^{-\frac{j}{2}} \psi(2^{-j} t - k), \quad j, k \in Z. \quad (4)$$

Transformata falkowa, w której funkcje bazowe są dyskretne, nazywana jest dyskretną transformatą falkową i definiowana jako:

$$W_f(j, k) = c_\psi^{-\frac{1}{2}} \int f(t) \psi_{j,k}^*(t) dt. \quad (5)$$

W przeciwieństwie do równania (2) uzyskanie transformaty odwrotnej do równania (5) nie jest łatwe. Konieczne staje się założenie o ortogonalności bazy falkowej $\{\psi_{j,k}(t), j, k \in \mathbb{Z}\}$. Wtedy transformata odwrotna do transformaty z równania (5) definiowana jest następująco:

$$f(t) = c_\psi^{-2} \sum_{j,k \in \mathbb{Z}} W_f(j,k) \psi_{j,k}(t). \quad (6)$$

Równanie (6) nazywane jest dyskretną falkową dekompozycją funkcji $f(t)$. W ogólnym przypadku sygnał $f(t)$ i składniki po prawej stronie równania (8) należą do wielowymiarowej przestrzeni $L^2(\mathbb{R}^n)$.

1.2. Sieć falkowo-neuronowa

Aproksymacja ciągłych funkcji ma fundamentalne znaczenie w zagadnieniach modelowania i identyfikacji systemów. Sztuczne sieci neuronowe uważane są za skuteczne narzędzie aproksymacji modeli o nieliniowych zależnościach między danymi wejściowymi a wyjściowymi. Innym podejściem do aproksymacji funkcji jest zastosowanie dyskretnej falkowej dekompozycji (8).

Funkcjami aktywacji neuronów sieci falkowo-neuronowej są funkcje falkowe (4), w ogólnym przypadku wielowymiarowe. Zaletą sieci falkowo-neuronowej jest to, iż istnieje bezpośredni związek między wartościami jej wag, które przypisane są połączeniom między neuronami a współczynnikami dyskretnej dekompozycji falkowej (6). Dzięki temu, w odróżnieniu do tradycyjnie sztucznej sieci neuronowej, gdzie inicjalizacja wartości wag sieci odbywa się najczęściej w sposób losowy, w sieci falkowo-neuronowej inicjalizacja wartości wag sieci odbywa się w sposób deterministyczny. Sieć falkowo-neuronową można opisać zależnością:

$$\hat{f}(t) = \sum_{i=1}^N w_i \psi(A_i(t - b_i)) + \bar{g}, \quad (7)$$

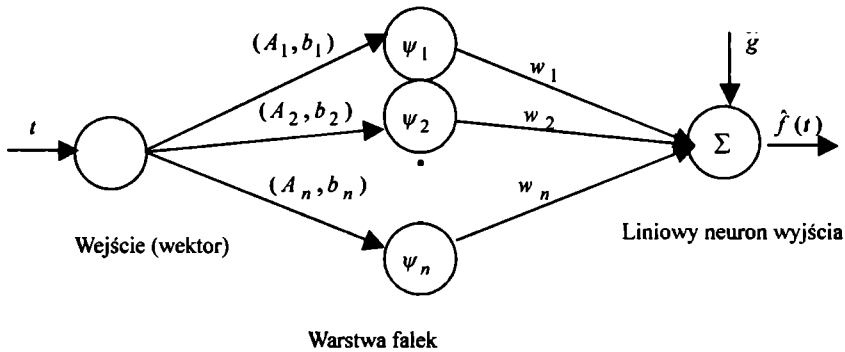
gdzie: $\psi(A_i(t - b_i))$ – falka wielowymiarowa,

b_i – wektor translacji,

A_i – macierz diagonalna zbudowana z wektorów odpowiedzialnych za skalowanie funkcji ψ ($A_i = \text{diag}(a_i)$, a_i – współczynnik skalowania),

$\bar{g}$ – parametr zwiększający efektywności sieci przy aproksymacji funkcji z niezerową wartością średnią (wartość średnia falki $\psi(t)$ jest zerem).

Oryginalny sygnał $f(t)$ z równania (6) jest aproksymowany przez liniową kombinację $\hat{f}(t)$ funkcji falkowych. Liniowe współczynniki w równaniu (7) odpowiadają wartościom wag sieci.



Rys. 1. Struktura sieci falkowo-neuronowej

Źródło: opracowanie własne.

Ogólna topologiczna struktura sieci falkowo-neuronowej została przedstawiona na rys. 1. Wartościami wag między neuronami w warstwie wejściowej sieci a jej warstwie ukrytej są współczynniki (A_i, b_i) , wartościom wag między neuronami w warstwie ukrytej a neuronem warstwy wyjściowej odpowiadają współczynniki w_i . Powyższe współczynniki podlegają optymalizacji poprzez zastosowanie algorytmu uczenia się sieci neuronowej (np. algorytmu opartego na wstecznej propagacji błędów).

2. Algorytm predykcji szeregu czasowego

Predykcja szeregu czasowego polega na oszacowaniu przyszłych wartości szeregu $x_{t+\tau}, x_{t+2\tau}, \dots$, jeśli ma się dane jego wartości z przeszłości $x_t, x_{t-\tau}, \dots, x_{t-(d-1)\tau}$. Wartości te tworzą tzw. przesuwalne okno czasowe (*sliding time window*). Istotą jest obserwacja, identyfikacja i modelowanie związku między wartościami przeszłymi a wartością przyszłą szeregu. Innymi słowy, należy znaleźć taką funkcję F , która spełnia odwzorowanie:

$$x_{t+\tau} = F(x_t, x_{t-\tau}, \dots, x_{t-(d-1)\tau}), \quad (8)$$

gdzie: d – wymiar okna czasowego,

τ_i – opóźnienia w czasie, $1 \leq i \leq d$.

Znalezienie dobrego odwzorowania lub aproksymacji dla nieznanej funkcji F może być dokonane przy użyciu kilku różnych technik aproksymacji. W tej pracy nazywaną techniką jest sieć falkowo-neuronowa, która została wyodrębniona z ogólnej teorii falkowej i sieci neuronowej. Konstrukcję tej sieci przedstawia poniższy algorytm:

1. Wartościami wejściowymi dla sieci są: skalarny szereg czasowy, wymiar d oraz opóźnienie τ .

2. Wybrana zostaje wielowymiarowa falka bazowa $\psi(A(t-b))$, ($A = \text{diag}(a)$, $\dim(A) = d$, gdzie a jest współczynnikiem skalowania) będąca funkcją aktywacji neuronów w warstwie ukrytej sieci.
3. Inicjowanie wartości wag sieci.
Do inicjowania wartości wag sieci wykorzystywany jest m.in. algorytm regresji. Spośród zbioru funkcji falkowych wybrany zostaje podzbiór, którego elementy minimalizują błąd regresji.

4. Wartości wag sieci zostają zoptymalizowane poprzez zastosowanie algorytmu uczenia sieci.
Aby wyznaczyć prognozę szeregu czasowego przy użyciu sieci falkowo-neuronowej, można zastosować co najmniej dwa podejścia. Jedno to przeprowadzenie prognozowania bezpośrednio na oryginalnym szeregu czasowym bez robienia redukcji zakłóceń, a drugie polega na prognozowaniu, gdy wcześniej usunięto już zakłócenia.

Dla ekonomicznych szeregów czasowych prostą metodą redukcji zakłóceń jest tzw. wygładzanie np. za pomocą średniej ruchomej. W związku z tym, w naszych rozważaniach w celu predykcji danego ekonomicznego szeregu czasowego zastosujemy średnią ruchomą, oczywiście jeśli wygładzanie szeregu będzie potrzebne. Dodatkowo, w celu aplikacji algorytmu służącego prognozowaniu, który byłby bardziej efektywny, usuniemy wykładnicze trendy wzrostu w każdym szeregu czasowym. Jest to konieczne, ponieważ opisywany algorytm jest bardziej dostosowany do stacjonarnych szeregów czasowych, podczas gdy ekonomiczny szereg czasowy z widocznym trendem jest niestacjonarny.

Jak wspominaliśmy wyżej, używamy dwóch metod do przeprowadzenia naszych testów: metoda z wygładzaniem szeregu czasowego i bez wygładzania.

Metoda z wygładzaniem

W tym przypadku kroki algorytmu są następujące:

1. Wygładzanie szeregu czasowego $x(1), x(2), \dots, x(N)$ za pomocą średniej ruchomej, co prowadzi do otrzymania nowego szeregu czasowego $y(i)$, $i = 12, 13, \dots, N$

$$y(i) = \frac{1}{12} (x(i-11) + x(i-10) + \dots + x(i-1) + x(i)). \quad (9)$$

2. Logarytmiczno-liniowa eliminacja trendu.

Aby otrzymać szereg czasowy $z(i)$, $i = 12, 13, \dots, N$ z wyeliminowanym trendem, stosujemy następującą relację:

$$z(i) = \ln y(i) - ai - b, \quad (10)$$

gdzie: parametry a , b mogą być otrzymane metodą najmniejszych kwadratów.

3. Jednokrokowa predykcja szeregu czasowego $z(i)$ przy użyciu sieci falkowo-neuronowej.
- Dzielimy zbiór danych na dwie części.
Pierwszą część danych, tzw. *zbiór treningowy*, używamy do dopasowania modelu predykcji, a drugą część, tzw. *zbiór testowy*, używamy do sprawdzenia wyniku predykcji, bazując na dopasowanym modelu.

- Wykorzystujemy wartości szeregu $z(i)$, $i = 12, 13, \dots, N_f (< N)$, aby nauczyć sieć falkowo-neuronową dopasowanego następującego modelu

$$z(i) = F(z(i - \tau_d), z(i - \tau_{d-1}), \dots, z(i - \tau_1)). \quad (11)$$

- Używamy pozostałe dane z szeregu $z(i)$, $i = N_f + 1, N_f + 2, \dots, N$, do sprawdzenia poprawności jednokrokowej predykcji, bazując na dopasowanym modelu.
4. Przesuwanie jednokrokowej predykcji od $z(i)$, $i = N_f + 1, N_f + 2, \dots, N$, z powrotem do predykcji odpowiednich $x(i)$'s. Faktycznie możemy najpierw przesunąć $z(i)$ do $y(i)$ według wzoru:

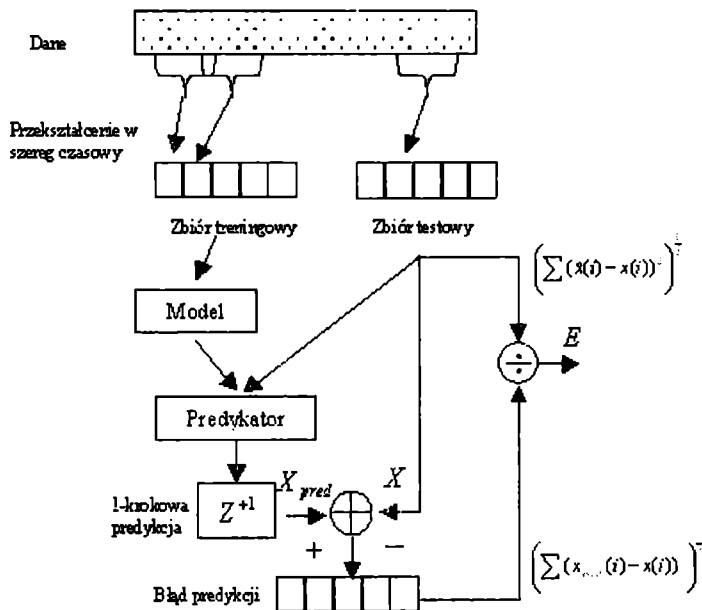
$$y(i) = e^{(z(i) + ai + b)}, \quad (12)$$

a następnie wykonać translację $y(i)$ do $x(i)$ według relacji:

$$x(i) = 12y(i) - x(i-11) - x(i-10) - \dots - x(i-1). \quad (13)$$

Metoda bez wygładzania

- Logarytmiczno-liniowa eliminacja trendu.
- Testowanie jednokrokowej predykcji szeregu czasowego $z(i)$ przy użyciu sieci falkowo-neuronowej.



Rys. 2. Schemat predykcji ekonomicznego szeregu czasowego

Źródło: opracowanie własne.

3. Przesuwanie jednokrokowej prognozy z $z(i)'s$ z powrotem do prognozy odpowiedniego $x(i)'s$ według wzoru:

$$x(i) = e^{(z(i)+ai+b)}. \quad (14)$$

2.1. Błąd predykcji

W celu oszacowania dokładności predykcji dokonujemy porównania między błędem predykcji generowanym przez sieć falkowo-neuronową a błędem generowanym przez predykcję polegającą na przyjęciu, że prognoza następnej wartości szeregu czasowego jest równa jego wartości przeszłej. Porównanie to określone jest wzorem

$$E = \left(\sum_i (x_{pred}(i) - x(i))^2 \right)^{\frac{1}{2}} / \left(\sum_i (x(i) - \hat{x}(i))^2 \right)^{\frac{1}{2}} \quad (15)$$

gdzie: $x(i)$ – dane z szeregu czasowego w chwili i ,

$x_{pred}(i)$ – predykcja wartości szeregu czasowego dokonana przez sieć falkowo-neuronową,

$\hat{x}(i) = x(i-1)$.

Jeżeli $E = 0$, to predykcja dokonana przez sieć falkowo-neuronową jest perfekcyjna, $E = 1$ oznacza, że prognoza dokonana przez sieć jest taka sama jak przez predyktor $\hat{x}(i)$.

3. Przykład numeryczny predykcji szeregu czasowego

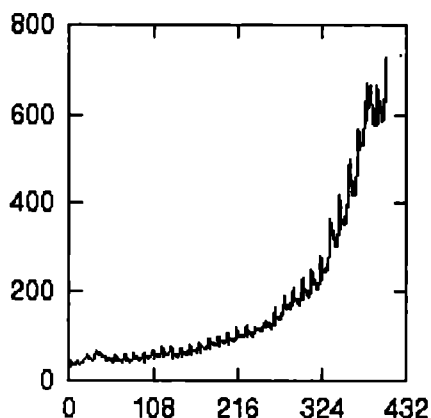
Dla celów prezentacji omówionego w punkcie 2 algorytmu predykcji szeregu czasowego przy użyciu sieci falkowo-neuronowej przyjmujemy na podstawie danych umownych ekonomiczny szereg czasowy (rys. 3) utworzony z 408 obserwacji, charakteryzujący się „dobrymi właściwościami” dla modelowanego zjawiska.

Postępując zgodnie z przedstawionym wcześniej algorytmem predykcji szeregu czasowego, wygładzamy rozpatrywany szereg za pomocą średniej ruchomej, a następnie dokonujemy logarytmiczno-liniowej eliminacji trendu, co ostatecznie prowadzi do otrzymania szeregu czasowego $z(i)$, $i = 12, 13, \dots, 408$, który przedstawiony jest na rys. 4.

Otrzymany wygładzony szereg czasowy dzielimy na dwa części. Pierwsza część, *zbiór treningowy*, zawiera 240 danych a część druga – *zbiór testowy*, 168 danych. Sieć poddaje się uczeniu dopasowanego modelu:

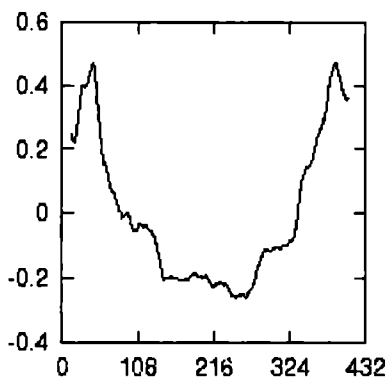
$$z(n) = F(z(n-1), z(n-6), z(n-12))$$

na grupie danych uczących zawartych w *zbiórze treningowym*, wykorzystując do tego zdolność sieci do zapamiętywania danych uczących.



Rys. 3. Wyjściowy szereg czasowy

Źródło: opracowanie własne.

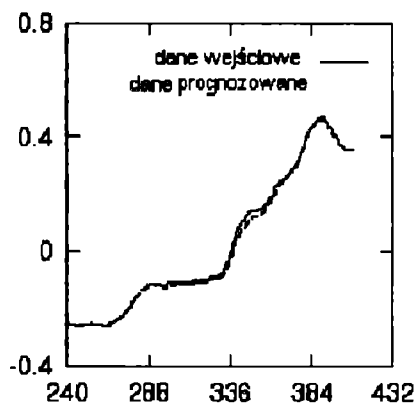


Rys. 4. Wygładzony szereg czasowy z wyeliminowanym trendem

Źródło: opracowanie własne.

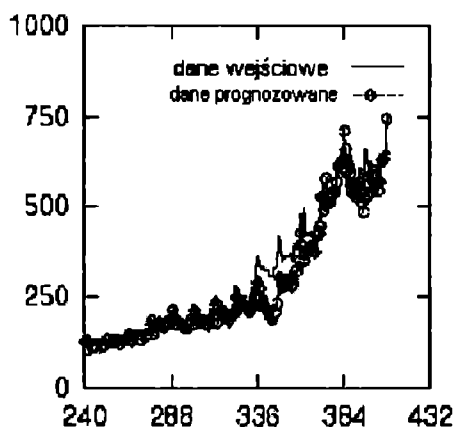
Zdolność rozpoznania danych *zbioru testowego* przez sieć, na których nie była uczona, jest miarą zdolności uogólniania. Efekt przeprowadzonej predykcji obrazuje rys. 5.

Jak obrazuje rys. 5, predykowane wartości dość dobrze odpowiadają wartościom rzeczywistym szeregu czasowego, co potwierdza obliczony na podstawie wzoru (15) błąd predykcji RMSE, który wynosi w tym przypadku 0,051. Gdy przesuniemy prognozowanie od $z(i)$'s z powrotem do prognozowania oryginalnych danych, to RMSE pomiędzy rzeczywistą i predygowaną wartością wyniesie 0,246. Reasumując, można stwierdzić, że prognozy na rozpatrywanym szeregu czasowym są zadowalające.



Rys. 5. Efekt jednokrokowej predykcji

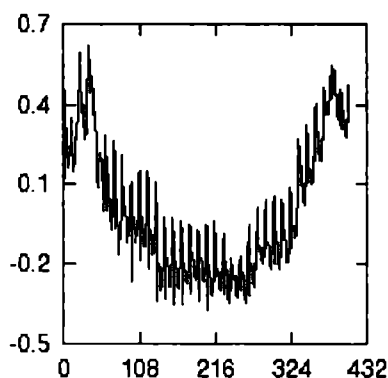
Źródło: opracowanie własne.



Rys. 6. Wynik przesunięcia predykcji wygładzonego szeregu z powrotem do prognozowania wyjściowych danych

Źródło: opracowanie własne.

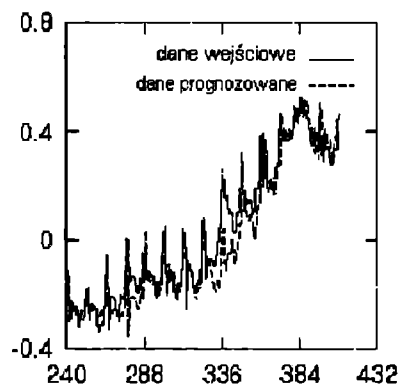
Aby zobaczyć efekt metody predykcji szeregu czasowego z wygładzaniem, zastosujemy teraz dla wyjściowego szeregu (rys. 3) metodę bez wygładzania. W tym przypadku omijamy wygładzanie szeregu przez średnią ruchomą, zaczynamy od logarytmiczno-liniowej eliminacji trendu czasowego, której wynik obrazuje rys. 7.



Rys. 7. Logarytmiczno-liniowa eliminacja trendu szeregu czasowego niewygładzonego

Źródło: opracowanie własne.

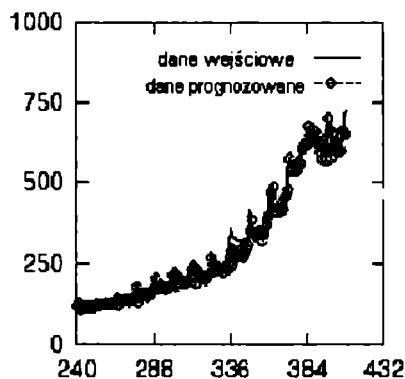
Tak jak wcześniej używamy pierwszych 240 danych szeregu czasowego z wyeliminowanym trendem do dopasowania odpowiedniego modelu predykcji, a drugą część do sprawdzenia predykcji. Rezultat predykcji niewygładzonego szeregu czasowego obrazuje rys. 8.



Rys. 8. Wynik jednokrokowej predykcji niewygładzonego szeregu czasowego

Źródło: opracowanie własne.

W tym przypadku błąd RMSE między rzeczywistą a prognozowaną wartością wynosi 0,172. Gdy przesuniemy prognozowanie od $z(i)$'s z powrotem do prognozowania oryginalnych danych, otrzymujemy:



Rys. 9. Rezultat przesunięcia predykcji niewygładzonego szeregu z powrotem do prognozowania wyjściowych danych

Źródło: opracowanie własne.

Rozpatrywany przykład predykcji szeregu czasowego pokazuje, że sieci falkowo-neuronowe mogą efektywnie prognozować ekonomiczne szeregi czasowe. Jednakże prognozy szeregu czasowego pierwszą metodą są o wiele lepsze niż metodą drugą, więc predykcja szeregu czasowego z wygładzaniem jest dokładniejsza niż bez wygładzania.

4. Zakończenie

W artykule zaproponowano metodę prognozowania ekonomicznych szeregów czasowych opartą na wykorzystaniu sieci falkowo-neuronowej. Wyniki wstępnych badań wskazują, że predykcja z wykorzystaniem sieci falkowo-neuronowej daje bardzo dobre wyniki. Wprawdzie jest obarczona niewielkim błędem, ale predyktor oparty na sieci falkowo-neuronowej jest lepszy od prostego predyktora opartego na wartości poprzedniej szeregu czasowego.

Reasumując możemy powiedzieć, że przykład wykazał, po przeprowadzeniu szeroko zakrojonych badań empirycznych, użyteczność sieci falkowo-neuronowych jako narzędzia krótkookresowego prognozowania zmiennych ekonomicznych.

Literatura

Aussem A., Murtagh F., *A neuro-wavelet Strategy for Web Traffic Forecasting*, „Journal of Official Statistics” 1998 no 1, s. 65-87.

- Cao L., Hong, Y., Fang, H., He G, *Predicting Chaotic Time Series with Wavelet Networks*, Physica D85 1995, s. 225-238.
- Mallat S.G., *Multiresolution Approximation and Wavelets Orthonormal Bases of $L^2(\mathbb{R})$* , Trans. Amer. Math Soc. 1989, no 1.
- Poggio T., Girosi F., *Networks for Approximation and Learning*. Proc, IEEE 1990, vol. 78, no 9.
- Yousefi'zadeh H., *Neural Network Estimation of Packet Arrival Rate in Self-Similar Queuing Systems*. Invited Paper. In *Proceedings of the Second Workshop on Fractals, Power Laws, and Other Next Generation Data Mining Tools*, ACM SIGKDD Conference 2003.
- Zhang Q., Benveniste A., *Wavelet Networks*, IEEE Transaction on Neural Networks 1992, vol. 3, s. 889-898.
- Zhang Q., *Using Wavelet Network in Nonparametric Estimation.*, IEEE Trans. Neural Networks 1997, vol. 8, s. 227-236.

USING WAVELET – NEURAL NETWORK TO PREDICTIVE ECONOMIC TIME SERIES

Summary

In this article a deterministic predictive technique is introduced which, is based on the embedding theorem by Takens and the recently developed wavelet networks. As a result, the predicted values economic time series correspond quite well with the actual values.