

**Paweł Lorek**

Akademia Ekonomiczna w Katowicach

---

## OBSZARY ZASTOSOWAŃ SIECI NEURONOWYCH W BIZNESIE

---

**Streszczenie:** W artykule opisano podstawowe zastosowania sieci neuronowych w biznesie. Dokonano syntetycznego zestawienia przykładowych zastosowań sztucznych sieci neuronowych w praktyce gospodarczej, ze szczególnym uwzględnieniem zalet i ograniczeń poszczególnych architektur sieci. Przedstawiono również koncepcję autorskiego systemu prognozowania cen miedzi, opartego na zespole sieci neuronowych. Działanie systemu zweryfikowano na danych pochodzących z notowań londyńskiej giełdy metali.

**Słowa kluczowe:** sieci neuronowe, zastosowania gospodarcze, prognoza szeregów czasowych, prognozowanie cen miedzi.

### 1. Wstęp

Sztuczne sieci neuronowe są metodami numerycznymi, zaliczanymi do grupy technik sztucznej inteligencji. Przez pół wieku swojego rozwoju znalazły liczne zastosowania w wielu obszarach nauki i techniki. Szczególną zaletą sieci neuronowych jest zdolność do uczenia się i adaptacji do zmieniających się warunków, co czyni je dobrym narzędziem analitycznym w przypadku zjawisk charakteryzujących się dużą dynamiką przebiegu.

W praktyce gospodarczej sieci neuronowe zyskują coraz mocniejszą pozycję i wciąż odkrywane są nowe możliwości ich zastosowania. Celem niniejszego artykułu jest dokonanie przeglądu znanych zastosowań sztucznych sieci neuronowych w biznesie oraz ukazanie skuteczności tych narzędzi na przykładzie prognozy światowych cen miedzi.

### 2. Sieci neuronowe a biznes

Powszechnym zjawiskiem w świecie biznesu jest powiększająca się złożoność procesów gospodarczych. Komplikacja tych procesów ma dwoistą naturę. Z jednej strony jest konsekwencją wciąż rozrastających się norm prawnych i przepisów. Z dru-

giej strony postępujący proces globalizacji wymusza uwzględnienie w planowaniu coraz większej liczby parametrów, presja zaś ze strony konkurencji skłania ku stosowaniu w procesie podejmowania decyzji różnego rodzaju prognoz i symulacji. O ile w pierwszym przypadku nie można liczyć na pomoc nauk ścisłych, o tyle w drugim powszechnie stosowane są różne techniki obliczeniowe, w tym metody sztucznej inteligencji.

Sztuczna inteligencja jest jednym ze szczytowych osiągnięć współczesnej informatyki. W jej obszarze znajdują się takie metody, jak [Rutkowski 2006] :

- sieci neuronowe (*neural networks*),
- logika rozmyta (*fuzzy logic*),
- algorytmy ewolucyjne (*evolutionary algorithms*),
- zbiory przybliżone (*rough sets*),
- metody probabilistyczne (*statistical methods*).

Trudno wyobrazić sobie współczesną gospodarkę bez udziału zdobywcy nauk informatycznych. Nie inaczej jest w wyżej wymienionych metodach, które nie tylko znajdują zastosowanie w biznesie w swojej czystej formie, ale występują również w postaciach hybrydowych [Bhargava, Gupta 2009] (np. rozmyte sieci neuronowe, ewolucyjna generacja reguł decyzyjnych).

Sieci neuronowe zajmują wśród metod sztucznej inteligencji miejsce szczególne, są bowiem szczytowym osiągnięciem tej dziedziny. Typowe klasy zastosowań sieci neuronowych w praktyce gospodarczej to [Osowski 2006] :

- klasyfikacja – sieć uczona jest rozpoznawania charakterystycznych cech wzorca, mających wpływ na ocenę badanego obiektu,
- predykcja – zagadnienie polegające na estymacji przyszłych wartości szeregu czasowego  $y_{n+1}, \dots, y_{n+k}$  na podstawie wartości przeszłych  $y_{n-m}, \dots, y_{n-1}$  oraz wartości obecnej  $y_n$ ,
- optymalizacja – dziedzina odpowiadająca programowaniu liniowemu i całkowitoliczbowemu, gdzie poszukuje się minimum zadanej funkcji celu, określonej przez uwarunkowania danego zadania.

Należy podkreślić, iż możliwości sieci neuronowych nie ograniczają się do wyżej wymienionych obszarów, jednakże w sferze zagadnień ekonomicznych, trudno wskazać problem obliczeniowy nie zaliczany do jednej z wymienionych kategorii. W zagadnieniach klasyfikacji najczęściej spotykanymi rodzajami sieci są: perceptron wielowarstwowy, mapa cech Kohonena i sieci rozmyte. Perceptron wielowarstwowy oraz sieci radialne są wykorzystywane w zagadnieniach predykcji. W zadaniach optymalizacji obok metod heurystycznych (algorytmy ewolucyjne, symulowane wyżarzanie, przeszukiwanie tabu) spotykane są sieci Hopfielda. Należy zdawać sobie sprawę, iż trudno przypisać rodzaj sieci do jednego konkretnego zadania. Normą jest sytuacja, w której jeden rodzaj sieci jest wykorzystywany w różnych aplikacjach, a nowe zastosowania wciąż są odkrywane. Przegląd zastosowań sieci neuronowych w poszczególnych sektorach gospodarki przedstawiono w tab. 1.

**Tabela 1.** Obszary zastosowań sieci neuronowych w biznesie

Sektor	Obszar zastosowań
Ubezpieczenia	wspomaganie decyzji i zarządzanie ryzykiem ubezpieczeniowym
Bankowość i finanse	ocena wiarygodności kredytobiorców, wykrywanie zjawiska prania brudnych pieniędzy, przewidywanie trendów giełdowych, przewidywanie kursów walut, ocena zagrożenia bankrutem, dywersyfikacja portfela inwestycyjnego, wspomaganie wyceny nieruchomości, ocena opłacalności inwestycji, weryfikacja opłacalności kontraktów
Logistyka	tworzenie harmonogramów przejazdów, planowanie tras przejazdu
Zarządzanie produkcją	kontrola procesu produkcji, zarządzanie projektami, planowanie operacji serwisowych, prognozowanie poziomu popytu
Marketing	identyfikacja charakterystyk konsumentów, prognozy sprzedaży, ukierunkowanie strategii marketingowych
Energetyka	prognozy zapotrzebowania na energię elektryczną, minimalizacja kosztów wytwarzania energii

Źródło: opracowanie własne na podstawie literatury.

Pomimo wielu możliwości użycia sieci neuronowych w biznesie, rzeczywistość stawia pewne bariery w ich stosowaniu. Do przeszkód można zaliczyć [Li 1994] :

- brak uniwersalnej metody doboru sieci do konkretnego zastosowania,
- zachowanie modelu opartego na sieci neuronowej może być nieprzewidywalne,
- duża dynamika zmian w procesach biznesowych wymusza częsty proces trenowania,
- śledzenie procesu uzyskiwania wyniku przez sieć neuronową jest bardzo trudne.

Pomimo wyżej wymienionych utrudnień można wskazać wiele przykładów, w których użycie sieci neuronowych zaowocowało otrzymaniem bardzo dobrych wyników. Jednym z oryginalniejszych zastosowań w sektorze finansów jest wykrywanie prania brudnych pieniędzy. System klasyfikacji został w tym wypadku oparty na samoorganizującej się mapie cech Kohonena. Zadaniem sieci jest wyodrębnienie z bazy danych operacji bankowych ciągów transakcji odpowiadających znanym schematom prania brudnych pieniędzy [Kuiljen, Migut 2004]. Mapa cech Kohonena ma również zastosowanie w segmentacji klientów. Jest to zadanie klasyfikacji polegające na wyodrębnieniu wśród klientów firmy podzbiorów o elementach charakteryzujących się zbliżonymi wartościami określonych cech [Migut 2004]. Podejście to pozwala lepiej uwzględnić potrzeby i oczekiwania klientów bez kosztownego personalizowania oferty dla każdego z klientów.

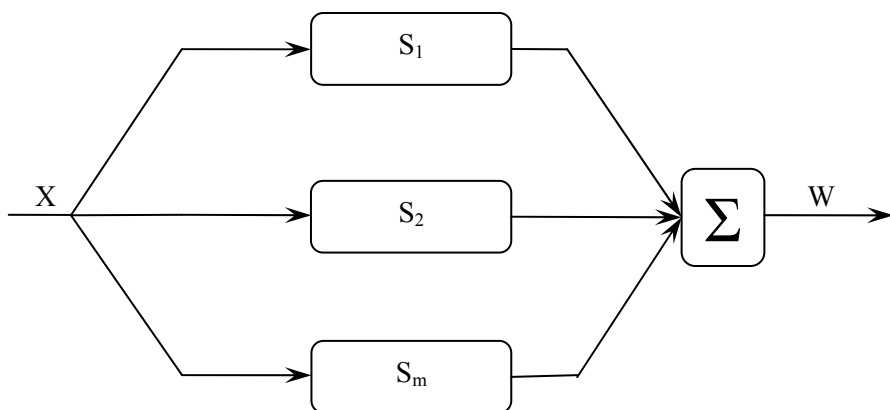
W świetle przyjętych przez Polskę umów międzynarodowych w zakresie ochrony środowiska coraz istotniejszym problemem staje się prognoza popytu na energię elektryczną oraz optymalizacja poziomu emisji zanieczyszczeń gazowych i pyłowych. W wypadku prognozy zapotrzebowania na energię elektryczną zagadnienie sprowadza się do estymacji szeregu czasowego. Spotykane są tutaj różne modele prognoz, mogące dotyczyć szacowania krótko-, średnio- i długoterminowego [Kash-

tiban, Tarafdar Haque 2005]. Najczęściej spotykane są jednak modele prognozy krótkoterminowej, w której najlepsze wyniki uzyskują systemy oparte na sieciach radialnych [Khan, Abraham 2009].

W sektorze energetycznym istotnym problemem jest również zaplanowanie odpowiedniego poziomu dostaw, gwarantującego minimalny koszt zakupu surowców energetycznych. Praktyka pokazuje zasadność użycia do tego celu sieci Hopfielda [Lee, Sode Yome 1998], aczkolwiek starsze modele nie uwzględniają ograniczeń środowiskowych, przyznaných limitów emisji gazów cieplarnianych oraz kosztu zakupu certyfikatów emisyjnych.

### 3. Zespół sieci neuronowych w predykcji światowych cen miedzi

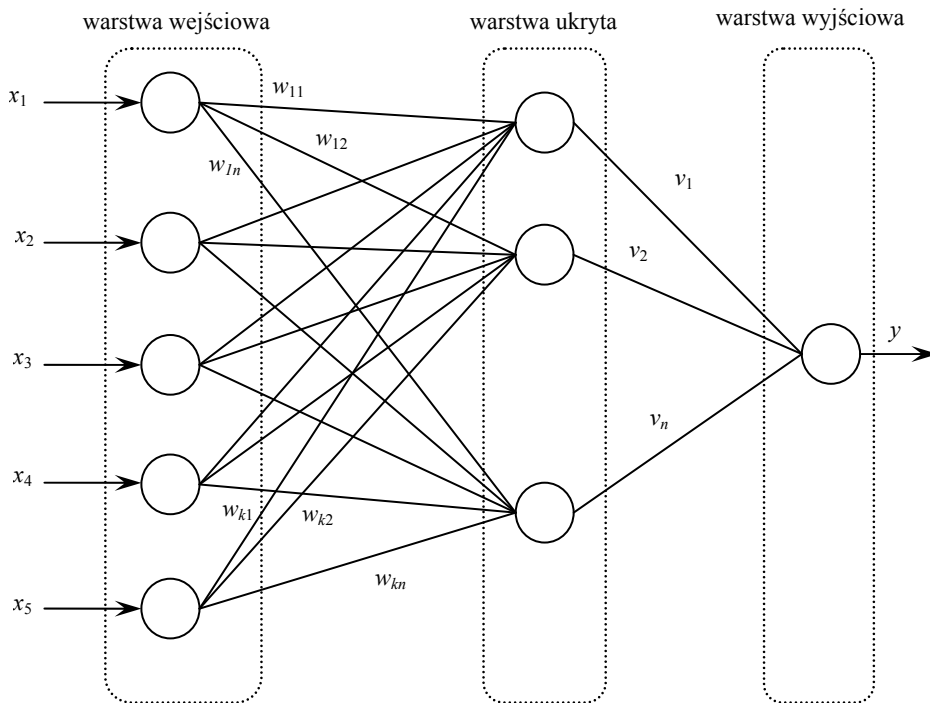
Jednym z trudniejszych zagadnień w sferze ekonomicznych zastosowań technik sztucznej inteligencji jest estymacja szeregów czasowych. Zasadniczą przeszkodą, na jaką można natrafić przy predykcji szeregów czasowych, jest ich stochastyczna natura. Narzędzie predykcji musi być na tyle dokładne, żeby wychwycić trend zmian, i na tyle ogólne, aby nie przejawiać wrażliwości na pojedyncze fluktuacje. Jako przykład takiego systemu został opracowany autorski układ predykcji cen miedzi na londyńskiej giełdzie metali (London Metal Exchange). Schemat systemu został przedstawiony na rys. 1.



Rys. 1. System predykcji za pomocą zespołu sieci neuronowych

Źródło : opracowanie własne.

Wektor  $X = \{x_t, x_{t+1}, x_{t+2}, x_{t+3}, x_{t+4}\}$  zawiera kolejne pięć wartości z badanego szeregu czasowego, z kolei wartość  $W$  jest estymacją wartości  $x_{t+5}$ . Rolę poszczególnych subestymatorów  $\{S_1, S_2, \dots, S_m\}$  pełnią sieci typu perceptron wielowarstwowy o strukturze przedstawionej na rys. 2.



Rys. 2. Sieć typu perceptron wielowarstwowy

Źródło: opracowanie własne.

Wyjście  $y$  sieci można opisać zależnością :

$$y = f \left( \sum_{i=1}^n v_i f \left( \sum_{j=1}^k w_{ji} x_j + w_{i0} \right) + v_0 \right),$$

- gdzie:
- wartości z prognozowanego szeregu czasowego,
  - $x_1, \dots, x_k$  – liczba neuronów warstwy ukrytej,
  - $w_{ji}$  – waga połączenia  $j$ -tego neuronu warstwy wejściowej z  $i$ -tym neuronem warstwy ukrytej,
  - $w_{i0}$  – próg aktywacji  $i$ -tego neuronu warstwy ukrytej,
  - $v_i$  – waga połączenia  $i$ -tego neuronu warstwy ukrytej z neuronem wyjściowym,
  - $v_0$  – próg aktywacji neuronu wyjściowego,
  - $f$  – funkcja aktywacji neuronu,  $f(x) = \tanh(x)$ .

Liczba neuronów w warstwie ukrytej (parametr  $n$ ) została ustalona zgodnie z twierdzeniem Kołmogorowa. Dla sieci o znormalizowanych wejściach o liczebności

$N = 5$ , wynosi ona  $2N+1=11$ . Dla tak zaprojektowanej sieci przygotowano zbiór uczący i testowy. Zbiór uczący składał się ze znormalizowanych dziennych cen miedzi z okresu styczeń 2006 – styczeń 2009, a testowy – ze znormalizowanych dziennych cen z okresu luty 2009 – maj 2009<sup>1</sup>. Jako metodę uczącą przyjęto algorytm propagacji wstecznej o współczynniku uczenia równym 0,1 i członie momentum równym 0,5. W badanym przypadku przyjęto liczbę przebiegów uczących równą 2000.

Sieci będące składowymi zespołu muszą spełniać kryteria dokładności i różnorodności. Zapewnienie odpowiedniej dokładności polega na określeniu górnego progu błędu testowania. Odpowiednią różnorodność można osiągnąć poprzez [Granitto, Verdes, Ceccato 2008]:

- użycie innego algorytmu uczenia dla każdej sieci,
- trenowanie poszczególnych sieci z różnymi wartościami współczynnika uczenia,
- użycie innych zestawów próbek uczących dla każdej sieci.

Ostatnie rozwiązanie jest wykorzystywane najczęściej. Dobór próbek do uczenia poszczególnych sieci jest dokonywany za pomocą metody *bagging* [Osowski 2006]. Algorytm ten polega na losowym doborze określonej liczby próbek z całego zbioru uczącego. Dzięki zastosowaniu tej metody każda sieć jest ćwiczona na zróżnicowanym zbiorze uczącym, co prowadzi do uzyskania różnych wyników dla poszczególnych sieci przy prezentacji tego samego zbioru testowego.

Należy zdawać sobie sprawę, iż dokładność i różnorodność są kryteriami wzajemnie się wykluczającymi. Trafny dobór sieci do zespołu predykcji polega więc na znalezieniu cienkiej granicy, na której te dwie cechy pozostają w równowadze.

Zespół sieci neuronowych został skonstruowany metodą superpozycji wyjść, za pomocą zależności:

$$W = \sum_{i=1}^k y_i \omega_i,$$

gdzie:  $W$  – wyjście zespołu sieci,

$k$  – liczba sieci w zespole,

$y_i$  – wyjście  $i$ -tej sieci w zespole,

$\omega_i$  – waga wyjścia  $i$ -tej sieci w zespole.

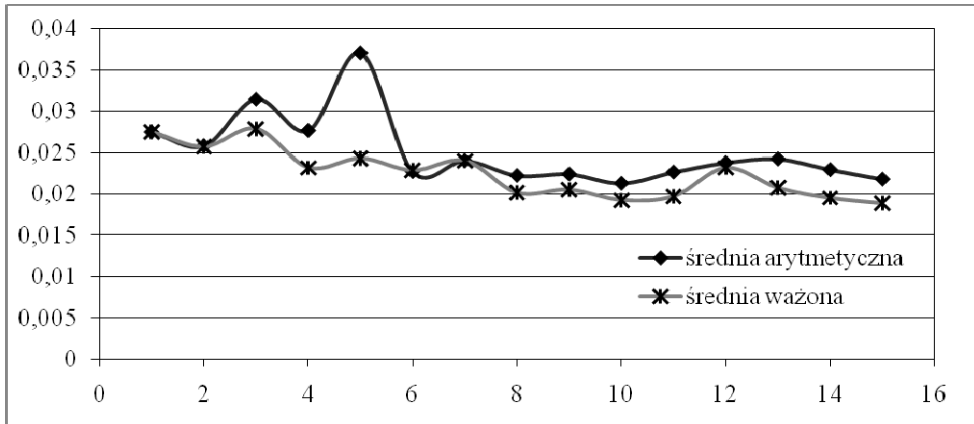
Wagi  $\omega_i$ , z jakimi sumowane są poszczególne wyjścia sieci, są wyznaczone w myśl zależności [Granitto, Verdes, Ceccato 2008]:

$$\omega_i = \frac{e_i^{-2}}{\sum_j e_j^{-2}},$$

gdzie:  $e_i, e_j$  – błąd średniokwadratowy na zbiorze testowym dla sieci  $i, j$  w zespole.

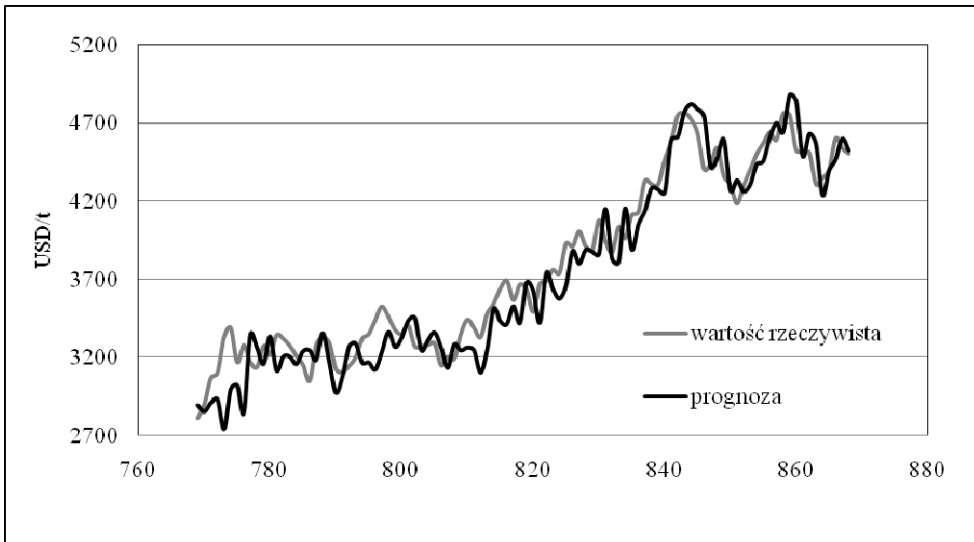
<sup>1</sup> Wykorzystane dane zostały pobrane ze strony [www.kghm.pl](http://www.kghm.pl), 1.06.2009.

Jako metodę alternatywną wypróbowano sumowanie wyjść poszczególnych sieci za pomocą średniej arytmetycznej. Wyniki porównania skuteczności działania obydwu metod przedstawiono na rys. 3.



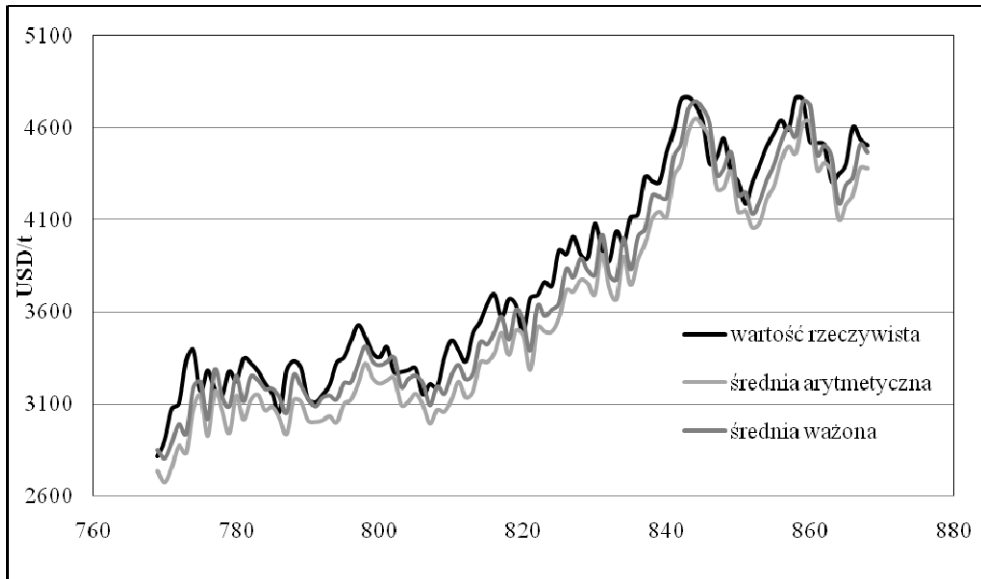
Rys. 3. Błąd średniokwadratowy predykcji na zbiorze testowym w funkcji liczby sieci w zespole

Źródło: opracowanie własne.



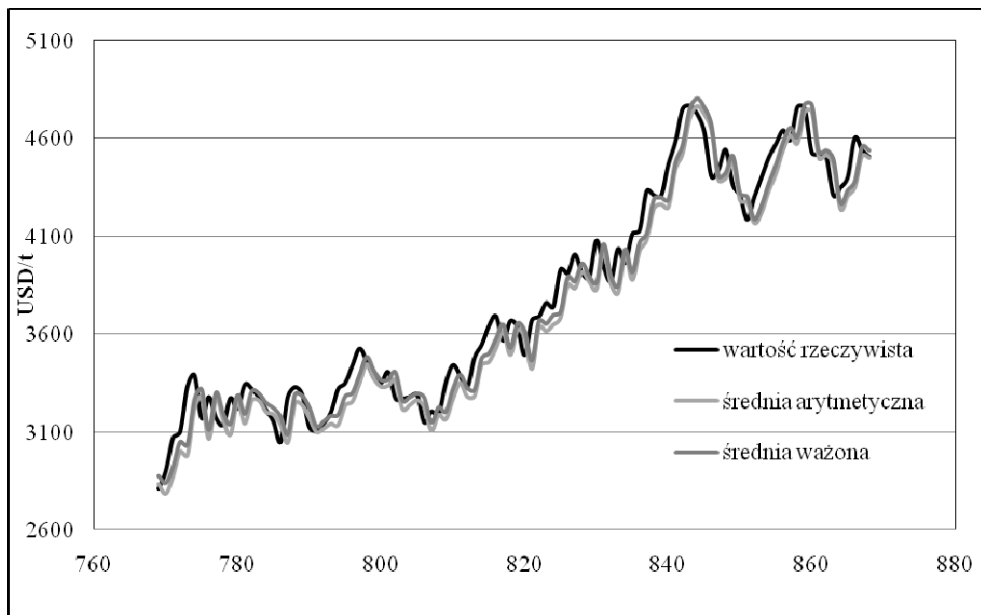
Rys. 4. Estymacja wartości zbioru testowego przez pojedynczą sieć

Źródło: opracowanie własne.



Rys. 5. Estymacja wartości zbioru testowego przez zespół 5 sieci

Źródło: opracowanie własne.



Rys. 6. Estymacja wartości zbioru testowego przez zespół 15 sieci

Źródło: opracowanie własne.



W przypadku obydwu metod zauważalny jest trend malejący przebiegu błędu średniokwadratowego predykcji w funkcji liczby sieci w zespole. Wpływ sieci o względnie dużym błędzie odwzorowania na jakość predykcji całego zespołu jest zdecydowanie silniejszy w przypadku stosowania średniej arytmetycznej. Zastosowana metoda średniej ważonej, przypisująca sieciom składowym wagi odwrotnie proporcjonalne do błędów średniokwadratowych odwzorowania, wykazuje takie zachowanie w dużo mniejszej skali. Ciekawym efektem obydwu metod jest coraz mniejsza różnica pomiędzy uzyskiwanymi wynikami w miarę wzrostu liczby sieci w zespole.

Wynik testowej predykcji, dokonanej przez pojedynczą sieć, przedstawiono na rys. 4. Zauważalna jest słaba dokładność prognozy w przypadku okresów charakteryzujących się brakiem bardzo wyraźnego trendu zmian cen.

Na rysunkach 5 i 6 przedstawiono wyniki testowej prognozy w przypadku zespołów sieci opartych na metodzie średniej arytmetycznej i ważonej, zawierających odpowiednio pięć i piętnaście sieci. Wyraźnie zauważalna jest tutaj poprawa jakości predykcji, jak również efekt uzyskiwania podobnych wyników w przypadku użycia różnych metod uśredniania dla zespołu o dużej liczbie sieci składowych.

#### 4. Wnioski i obserwacje

Przeprowadzone doświadczenia pozwalają sformułować następujące wnioski i obserwacje:

- jakość predykcji szeregu czasowego może być poprawiona poprzez zastosowanie zespołu sieci neuronowych zamiast pojedynczej sieci,
- stopień poprawy zależy od cech estymowanego szeregu czasowego i nie może być określony *a priori*,
- średnia ważona oparta na błędzie odwzorowania poszczególnych sieci daje lepsze wyniki niż uśrednianie wyjść za pomocą średniej arytmetycznej,
- uśrednianie wyjść poszczególnych sieci za pomocą średniej arytmetycznej wykazuje większą wrażliwość na obecność w zespole sieci o gorszej jakości odwzorowania.

#### 5. Podsumowanie

Przeprowadzone doświadczenia, jak i literatura przedmiotu pokazują, że sztuczne sieci neuronowe mogą być z powodzeniem stosowane jako instrument wspomagania decyzji w biznesie. O popularności tych metod świadczą liczne zastosowania w różnych dziedzinach gospodarki. Przy stosowaniu wymienionych technik należy jednak zachować daleko posuniętą ostrożność. Zachowanie modeli opartych na sieciach neuronowych może być nieprzewidywalne pomimo obiecujących rezultatów uzyskanych w trakcie testowania. Jest to jeden z zarzutów wobec stosowania wspomnianych metod. Do innych zarzutów można zaliczyć brak wartości poznawczej. Istot-

nie, sposób działania sieci neuronowej praktycznie wyklucza możliwość poznania natury badanego zjawiska. Również śledzenie sposobu otrzymania rozwiązania przez sieć jest bardzo utrudnione. Należy jednak spojrzeć na tę kwestię również z innego punktu. Proces selekcji danych uczących i testowych, doboru rodzaju i architektury sieci, algorytmu i parametrów uczenia czy wreszcie interpretacja i opis uzyskanych wyników – pozwalają uzyskać cenną wiedzę o stosowanych algorytmach i doświadczenie. Nie ma żadnych podstaw, aby oceniać tę wiedzę jako mniej wartościową lub mniej znaczącą.

## Literatura

- Bhargava N., Gupta M., *Application of artificial neural networks in business applications*, [www.cse.iitd.ernet.in/~mcs042650/lahore\\_conference.pdf](http://www.cse.iitd.ernet.in/~mcs042650/lahore_conference.pdf), 15.05.2009.
- Granitto P.M., Verdes P.F., Ceccato H.A., *Neural Networks Ensembles: Evaluation of Aggregation Algorithms*, [arxiv.org/abs/cs/0502006v1](http://arxiv.org/abs/cs/0502006v1), 9.05.2009.
- Kashtiban A.M., Tarafdar Haque M., *Application of neural networks in power systems; A review*, [www.waset.org/pwaset/v6/v6-12.pdf](http://www.waset.org/pwaset/v6/v6-12.pdf), 23.04.2009.
- Khan M., Abraham M., *Short term load forecasting on Czech Republic using soft computing paradigms*, [arxiv.org/abs/cs.AI/0405051](http://arxiv.org/abs/cs.AI/0405051), 5.05.2009.
- Kuiljen T., Migut G., *Wykrywanie nadużyć i prania brudnych pieniędzy*, [www.statsoft.pl](http://www.statsoft.pl), 10.05.2009.
- Lee K.Y., Sode-Yome A., *Adaptive Hopfield neural networks for economic load dispatch*, [labs.ee.psu.edu/labs/powerlab/papers/power\\_9805.pdf](http://labs.ee.psu.edu/labs/powerlab/papers/power_9805.pdf), 11.04.2009.
- Li E.Y., *Artificial neural networks and their business applications*, [www.cob.calpoly.edu/~eli/pdf/neural.pdf](http://www.cob.calpoly.edu/~eli/pdf/neural.pdf), 6.05.2009.
- Migut G., *Jak znaleźć grupy podobnych klientów czyli metody segmentacji*, [www.statsoft.pl](http://www.statsoft.pl), 10.05.2009.
- Rutkowski L., *Metody i techniki sztucznej inteligencji*, Wydawnictwo Naukowe PWN, Warszawa 2006.
- Osowski S., *Sieci neuronowe do przetwarzania informacji*, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2006.

## NEURAL NETWORKS IN BUSINESS APPLICATIONS

**Summary:** The article discusses basic applications of neural networks in business with the focus on their advantages and disadvantages. The description of system based on the set of neural networks is presented in the second part of the article. The main purpose of this system is forecasting the copper prices.