INFORMATYKA EKONOMICZNA BUSINESS INFORMATICS 22 • 2011

ISSN 1507-3858

Paweł Lorek
Uniwersytet Ekonomiczny w Katowicach

PRZYDATNOŚĆ WYBRANYCH TYPÓW SIECI NEURONOWYCH DO PROGNOZOWANIA ZMIENNOŚCI CEN ENERGII ELEKTRYCZNEJ

Streszczenie: W artykule poruszono problem prognozowania zmienności dziennych cen energii na giełdzie Nord Pool Spot. Zaproponowano dwa alternatywne modele bazujące na sztucznych sieciach neuronowych – perceptronie wielowarstwowym oraz sieci realizującej regresję uogólnioną. Szczególny nacisk położono na porównanie dokładności prognozowania obydwu modeli oraz odmienność ich zachowania.

Słowa kluczowe: sieci neuronowe, zastosowania gospodarcze, prognoza szeregów czasowych, giełdy energii.

1. Wstęp

Postępujący proces liberalizacji różnych sektorów gospodarki nie ominął sektora elektroenergetycznego. Sektor ten przez lata był uważany za obszar monopolu naturalnego. Istniało również wiele obiekcji natury politycznej, koncentrujących się wokół argumentacji o strategicznym znaczeniu. Nic więc dziwnego, że obrót energią elektryczną ulegał procesowi liberalizacji stosunkowo długo, poprzez wiele form pośrednich między gospodarką regulowaną a rynkową. Należy podkreślić, że pełna liberalizacja, głównie za sprawą uwarunkowań technicznych, nie jest w tym przypadku możliwa. Jednym z symptomów liberalizacji jest proces powstawania i rozwoju giełd energii elektrycznej. Rynek giełdowy jest powszechnie uważany za formę organizacyjną gwarantującą najlepszą alokację zasobów towarowych i pieniężnych. Forma ta została zaadaptowana również na potrzeby sektora elektroenergetycznego, pomimo istnienia wielu ograniczeń natury technicznej, charakterystycznych dla tego segmentu gospodarki. Naturalną konsekwencją pojawienia się nowych rynkowych form organizacji w sektorach dotychczas regulowanych jest weryfikacja przydatności instrumentów zarządzania ryzykiem, używanych dotychczas na innych rynkach. Celem artykułu jest sprawdzenie przydatności sieci neuronowych typu perceptron wielowarstwowy oraz sieci realizującej regresję uogólnio-

ną jako narzędzi prognostycznych i zweryfikowanie tej przydatności na przykładzie prognozy notowań wybranej europejskiej giełdy energii.

Jednym z najczęściej badanych zastosowań sieci neuronowych do zadań prognostycznych jest prognozowanie na rynkach kapitałowych. Udane próby były odnotowywane w przypadku rynków zagranicznych [Li 1994, s. 7], jak również rynku krajowego [Timofiejczuk 2004, s. 111]. W przypadku prognoz na rynku finansowym najlepsze rezultaty przynosi użycie perceptronu wielowarstwowego [Timofiejczuk 2004, s. 165]. Oprócz prognozowania na rynkach kapitałowych sieci neuronowe są spotykane jako narzędzia prognostyczne na rynkach towarowych. Wśród udanych zastosowań można wskazać system do predykcji światowych cen ropy naftowej. Podobnie jak w poprzednim przypadku, często spotyka się tu sieć typu perceptron wielowarstwowy, ale badane są też sieci rekurencyjne (np. sieć Elmana) [Kulkarni, Haidar 2009, s. 5].

Sieci neuronowe znajdują również zastosowanie do prognozowania wskaźników makroekonomicznych. Przykładem jest system prognozujący wskaźnik IGAEM (*Monthly Economic Activity General Indicator*). Wskaźnik ten jest wyliczany przez Bank Centralny Wenezueli i służy jako uzupełnienie wskaźnika PKB. Symulacja tak złożonego systemu jak gospodarka państwa wymagała użycia sieci neuronowej o niestandardowej, bardziej skomplikowanej architekturze. W tej roli użyto hierarchicznego systemu złożonego z kilku sieci typu perceptron wielowarstwowy o jednej warstwie ukrytej [Constanzo i in. 2010, s. 10].

2. Dane źródłowe

Jako dane uczące zostały wykorzystane dzienne ceny energii elektrycznej na rynku Nord Pool Spot, pochodzące z okresu 1.01.2007–31.12.2010. W charakterze danych testowych wykorzystano notowania z okresu 1.01.2011–30.06.2011[1].

Wstępna obróbka danych, oprócz podziału na zbiór uczący i testowy, obejmowała przekształcenie szeregów cen w szeregi logarytmicznych stóp zwrotu, zgodnie z zależnością [Jajuga 2011, s. 7]:

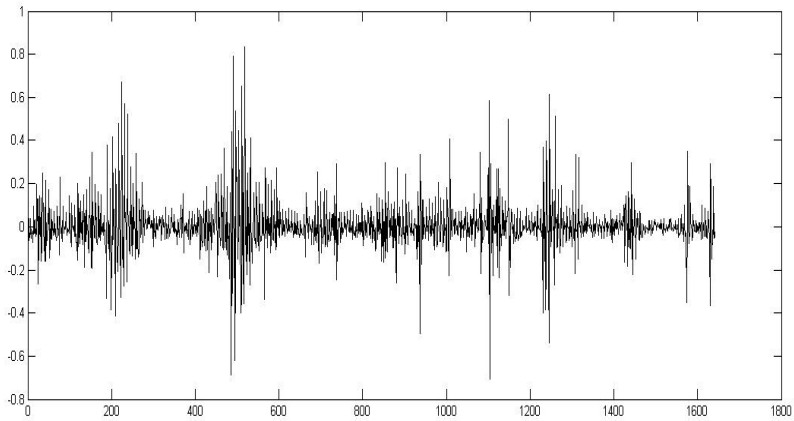
$$R = \ln \frac{P_{t+1}}{P_t},$$

gdzie: P_t – wartość szeregu w momencie t,

P_{t+1} – wartość szeregu w momencie t+1.

Łączne zestawienie przekształconych szeregów przedstawia rys. 1.

[1] Wszystkie wykorzystane dane zostały pobrane ze strony www.nordpoolspot.com.



Rys. 1. Szereg logarytmicznych stóp zwrotu

Źródło: opracowanie własne.

Wykonanie przekształcenia cen na logarytmiczne stopy zwrotów pozwala na eliminację trendu z danych uczących i testowych, co jest warunkiem koniecznym przy przygotowaniu danych do uczenia i testowania sieci neuronowych.

Jako wskaźnik jakości prognozy został przyjęty błąd średniokwadratowy zdefiniowany następująco:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

gdzie: n – długość zbioru testowego,

y_i – wartość elementu zbioru testowego,

$\hat{y}_i$ – prognozowana wartość elementu zbioru testowego.

Błąd średniokwadratowy MSE jest zdecydowanie najczęściej wybieranym wskaźnikiem jakości wykonanych prognoz. Jego ilościowa postać pozwala na porównanie jakości opracowanego modelu z innymi.

3. Wykorzystanie wybranych sieci neuronowych do celów prognostycznych

Perceptron wielowarstwowy (*Multi Layer Perceptron* – MLP) jest architekturą najczęściej spotykaną w zastosowaniach praktycznych. Jest opisany w wielu monografiach poświęconych sztucznej inteligencji [Osowski 2006, s. 44; Rutkowski 2006, s. 180]. Problem zaprojektowania adekwatnej sieci perceptronowej do zadań prognostycznych sprowadza się do ustalenia liczby neuronów wejściowych oraz

liczby neuronów ukrytych, w czym pomocna jest obserwacja błędu działania sieci na wyselekcjonowanym zbiorze testowym.

Do podstawowych cech sieci MLP zaliczyć można [Tadeusiewicz, Lula 2000, s. 539]:

- stosunkowo długi czas trenowania,
- dobre właściwości ekstrapolacyjne,
- warstwowa architektura (najczęściej trójwarstwowa),
- konieczność skalowania danych wejściowych ze względu na specyfikę funkcji aktywacji neuronów.

Jedną z istotniejszych cech sieci MLP jest również fakt, że celowe i skuteczne działanie wykazują sieci nawet o bardzo ubogiej strukturze.

Sieć realizująca regresję uogólnioną (*Generalized Regression Neural Network* – GRNN) została opracowana przez D. Spechta w 1991 r. W stosunku do sieci MLP jest ona zdecydowanie mniej popularna. Znajduje to również odzwierciedlenie w literaturze, która na temat tej sieci jest stosunkowo uboga. Zasadne więc jest skrótowe opisanie zasady jej działania.

Sieci GRNN zawsze składają się z czterech warstw: wejściowej, radialnej, regresyjnej oraz wyjściowej [Tadeusiewicz 2007, s. 67]. Neurony w warstwie wejściowej mają za zadanie, podobnie jak w sieci MLP, przechowywanie wartości wejściowych. W warstwie radialnej znajdują się neurony o gaussowskich funkcjach aktywacji, dla których największa wartość wyjściowa jest osiągana dla poszczególnych elementów uczących. Liczebność tej warstwy zawsze jest równa liczbie próbek w zbiorze uczącym. W warstwie regresyjnej dokonywane jest obliczenie sumy wag oraz sumy wartości wyjściowych wchodzących w skład zbiory uczącego. W warstwie wyjściowej obliczany jest iloraz wartości obliczanych w warstwie regresyjnej, który stanowi wartość wyjściową sieci [Tadeusiewicz, Lula 2000, s. 541].

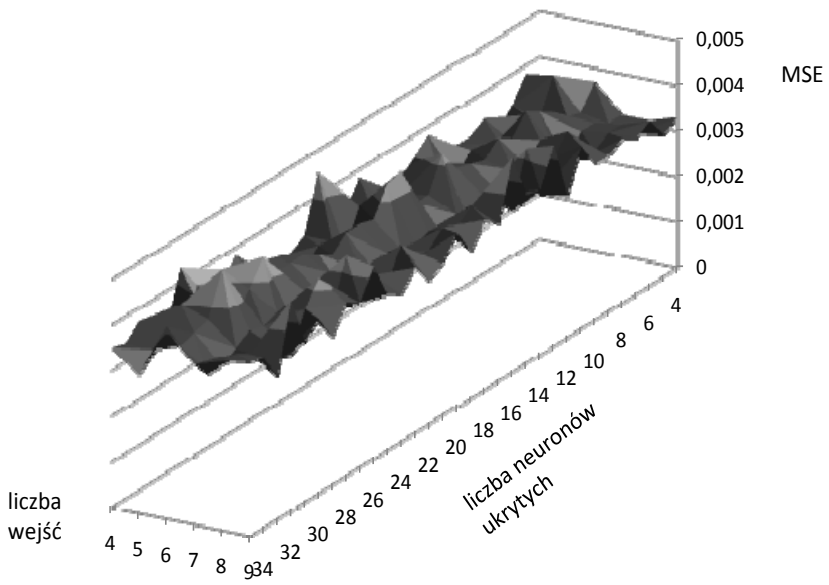
Próbując syntetycznie ująć charakterystyczne cechy sieci GRNN, można dokonać następującego zestawienia [Tadeusiewicz, Lula 2000, s. 542]:

- krótki czas trenowania,
- słabe zdolności ekstrapolacyjne,
- warstwowa architektura (zawsze czterowarstwowa),
- bardzo duże rozmiary wytrenowanych sieci.

Jak wynika z powyższego opisu, sieć typu GRNN charakteryzuje odmienna filozofia działania. Cecha ta powinna zadecydować o odmiennej charakterystyce uzyskanych wyników prognoz przez ten typ sieci. Proces projektowania sieci typu GRNN jest jednak podobny jak w sieci MLP. Podobnie jak w poprzednim przypadku, należy na drodze eksperymentalnej ustalić liczbę neuronów wejściowych. Liczba neuronów ukrytych jest zdeterminowana liczbą elementów w zbiorze uczącym, ale potrzebne jest ustalenie parametrów σ gaussowskich funkcjach aktywacji tych neuronów.

4. Weryfikacja modelu opartego na sieci MLP

Wpływ rozmiaru wektora wejściowego oraz liczby neuronów w warstwie ukrytej na wartość błędu testowania sieci MLP zaprezentowano na rys. 2. Na przedstawionym wykresie odznacza się wyraźnie spadek błędu prognozy w przypadku wykorzystania sieci o siedmiu neuronach wejściowych. Prawdopodobnym wyjaśnieniem tego zjawiska wydaje się obecność zmian cyklicznych o okresie równym siedmiu dniom.



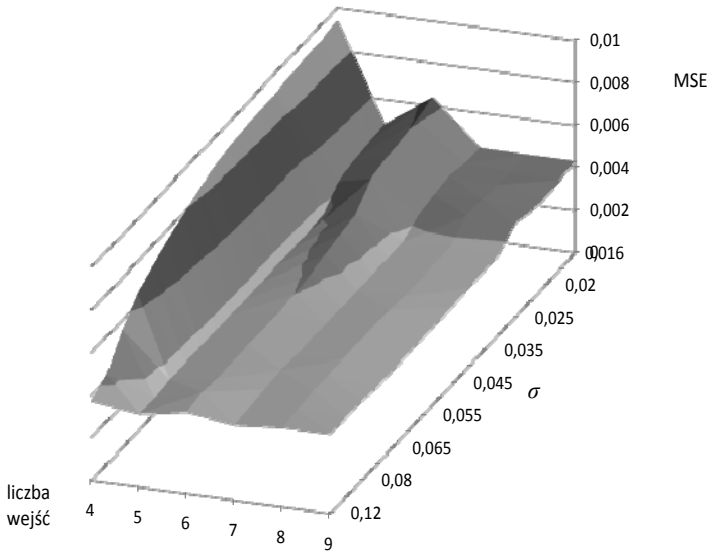
Rys. 2. Powierzchnia błędu testowania dla sieci MLP

Źródło: opracowanie własne.

Godnym uwagi zjawiskiem jest wspomniane wcześniej celowe i (względnie) skuteczne działanie nawet bardzo prostych sieci. Zachowanie takie wykazują sieci nawet o zaledwie czterech neuronach w warstwie ukrytej. Kolejnym potwierdzeniem teoretycznych cech perceptronu jest brak poprawy działania przy powiększaniu rozmiarów warstwy ukrytej ponad określony pułap. Z przedstawionego wykresu wynika, że sieć wykazuje dużo większą wrażliwość na zmiany liczby neuronów w warstwie wejściowej niż w warstwie ukrytej. Minimalnym błędem wśród wszystkich badanych sieci MLP wykazała się w tym przypadku sieć o 7 neuronach wejściowych i 15 neuronach ukrytych, osiągając błąd testowania MSE = 0,0027.

5. Weryfikacja modelu opartego na sieci GRNN

Na rysunku 3 przedstawiona została powierzchnia błędu testowania sieci GRNN. Zwracającym uwagę charakterystycznym elementem jest podwójne minimum funkcji błędu testowania występujące dla 5 oraz 7 neuronów w warstwie wejściowej, przy czym minimum dla 7 neuronów wejściowych ma charakter minimum lokalnego, natomiast dla 5 neuronów wejściowych minimum globalnego. Zachowanie sieci GRNN jest w tym przypadku inne niż sieci MLP, aczkolwiek występowanie minimum dla 7 neuronów wejściowych świadczy o pewnej dozie podobieństwa. Kolejnym podobieństwem jest duży wpływ liczby neuronów wejściowych na jakość prognozy wykonywanej przez sieć.



Rys. 3. Powierzchnia błędu testowania dla sieci GRNN

Źródło: opracowanie własne.

W przeprowadzonych obliczeniach zwraca uwagę podobny, ogólny poziom błędu testowania dla obu badanych sieci, aczkolwiek minimalne błędy prognozy dla sieci MLP (MSE = 0,0027) oraz GRNN (MSE = 0,0035) skłaniają do wysunięcia wniosku o lekkiej przewadze sieci MLP. Istotną cechą jest również dużo mniejsza wrażliwość sieci GRNN na zmiany parametrów topologicznych, o czym świadczy gładsza struktura powierzchni błędu. Z przedstawionego przebiegu funkcji błędu oraz opisanej specyfiki sieci GRNN wynika potwierdzenie jej spodziewanego odmiennego zachowania w stosunku do sieci MLP.

6. Podsumowanie

Literatura przedmiotu przedstawia wiele udanych zastosowań sieci neuronowych do zadań prognostycznych. Kolejnym dowodem na ich przydatność są przedstawione w niniejszym artykule wyniki obliczeń. Prowadzą one do następujących wniosków:

- zarówno w przypadku sieci typu MLP, jak i GRNN jakość prognozy zależna jest od określonej struktury sieci,
- w przypadku obydwu typów sieci parametrem mającym decydujący wpływ na jakość prognozy jest liczba neuronów w warstwie wejściowej,
- obydwie sieci wykazały zbliżony poziom błędu na zbiorze testowym, chociaż sieć typu MLP wykazała się działaniem minimalnie lepszym,
- różnice w działaniu obydwu typów sieci uwidocznione są poprzez wygląd uzyskanych powierzchni błędu, która jest zdecydowanie gładsza w przypadku sieci typu GRNN.

Wnioski te potwierdzają tezę głoszącą, że prognozowanie ekonomicznych szeregów czasowych jest zagadnieniem złożonym. Idealną sytuacją jest, gdy uda się zidentyfikować czynniki wpływające na dane zjawisko i na ich bazie skonstruować model bezpośrednio ekstrapolujący odpowiednie wartości. Zdarza się to jednak niezwykle rzadko, gdyż złożoność procesów gospodarczych okazuje się fundamentalną przeszkodą. Należy wówczas sięgnąć po metody oparte na zgromadzonych obserwacjach badanego zjawiska. Do takich metod zaliczają się m.in. sieci neuronowe, użyte w doświadczeniach opisanych w niniejszym artykule.

Literatura

Constanzo S., Trigo L., Jiménez L., González J. [2010], *A Neural Networks Model of the Venezuelan Economy*, arxiv.org/abs/0708.3463 [15.06.2011].

Jajuga K. [2011], *Metody statystyczne w finansach*, www.statsoft.pl [23.06.2011].

Kulkarni S., Haidar I. [2009], *Forecasting Model for Crude Oil Price Using Artificial Neural Networks and Commodity Futures Prices*, arxiv.org/abs/0906.4838 [19.06.2011].

Li E. [1994], *Artificial Neural Networks and Their Business Applications*, www.cob.calpoly.edu/~eli/pdf/neural.pdf [25.06.2011].

Osowski S. [2006], *Sieci neuronowe do przetwarzania informacji*, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa.

Rutkowski L. [2006], *Metody i techniki sztucznej inteligencji*, Wydawnictwo Naukowe PWN, Warszawa.

Tadeusiewicz R. [2007], *Odkrywanie właściwości sieci neuronowych przy użyciu programów w języku C#*, Polska Akademia Umiejętności, Kraków.

Tadeusiewicz R., Lula P. [2000], *Neuronowe metody analizy szeregów czasowych i możliwości ich zastosowań w zagadnieniach biomedycznych*, [w:] *Biocybernetyka i inżynieria biomedyczna*, t. 6:

Sieci neuronowe, red. W. Duch, J. Korbicz, L. Rutkowski, R. Tadeusiewicz, Akademicka Oficyna Wydawnicza Exit, Warszawa.

Timofiejczuk G. [2004], *Zastosowanie sieci neuronowych do prognozowania na rynkach finansowych, praca doktorska*, AE Katowice.

USEFULNESS OF SELECTED TYPES OF NEURAL NETWORKS FOR FORECASTING OF ELECTRICITY PRICE VOLATILITY

Summary: The subject of the article is forecasting of daily changes of Nord Pool Spot system price. Two ways of forecasting are discussed. The first is based on the multilayer perceptron, the second on the generalized regression neural network. The results of applying both methods are compared and discussed.

Keywords: neural networks, time series forecasting, energy exchange.