

Sławomir Francik
Katedra Inżynierii Mechanicznej i Agrofizyki
Akademia Rolnicza w Krakowie

PROGNOZOWANIE CENY OGÓRKA SZKLARNIOWEGO ZA POMOCĄ SIECI NEURONOWYCH

Streszczenie

W pracy opracowano modele wykorzystujące sztuczne sieci neuronowe do prognozowania cen ogórka szklarniowego, przy czterech horyzontach prognoz. Porównano dokładności prognoz uzyskanych za pomocą różnych typów sieci neuronowych (liniowych, wielowarstwowych perceptronów i sieci o radialnych funkcjach bazowych). Jako najlepsze modele wybrano sieci liniowe, gdyż pozwalały na uzyskanie najdokładniejszych prognoz.

Słowa kluczowe: sztuczne sieci neuronowe, prognozowanie.

Wprowadzenie

Analizując zmiany cen produktów na giełdach rolno - spożywczych można zauważyć, że ceny te zmieniają się w gwałtownie w ciągu roku – często występuje okresowa zmienność cen. A zatem moment sprzedaży przez rolnika swojego produktu w istotny sposób wpływa na dochodowość prowadzonego przez niego gospodarstwa. Właściwa decyzja uwzględniająca prognozy zmian cen, koszty przechowywania i ewentualnych strat plonu, decyduje o uzyskanym zysku. Większe zyski umożliwiają większe inwestycje związane z prowadzoną produkcją rolniczą i rozwój gospodarstwa. Na przykład cena zbytu 1kg ogórka szklarniowego zmieniała się w latach 2001 – 2004 od około 0,50 [zł/kg] do ponad 10 [zł/kg] – a zatem maksymalna cena jest dwudziestokrotnie większa od ceny minimalnej.

Jedną z informacji pozwalających zoptymalizować decyzje o sprzedaży jest prognoza zmian cen. Zalety sztucznych sieci neuronowych wskazują na możliwość ich zastosowania do prognozowania cen na giełdach rolno spożywczych [1, 2, 3].

Cel pracy i metodyka

W pracy podjęto próbę opracowania modelu zmiany cen ogórka szklarniowego. Do tworzenia modelu zastosowano sztuczne sieci neuronowe (SSN). Jako dane historyczne, konieczne do uczenia i testowania sztucznych sieci neuronowych, posłużyły średnie ceny ogórka szklarniowego na giełdzie rolno – spożywczej od 1.09.2000 do 17.12.2004. Aby uzyskać cenę dla każdego dnia w roku, w dniach pomiędzy notowaniami cenę ustalono stosując interpolację liniową. Sformułowano cztery modele prognostyczne, dla różnych horyzontów czasowych prognozy (od 1 do 4 tygodni):

$$\text{model 1: } y_1 = F(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$$

$$\text{model 2: } y_2 = F(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$$

$$\text{model 3: } y_3 = F(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$$

$$\text{model 4: } y_4 = F(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$$

gdzie:

- x_1 – dzień prognozy T_0 ,
- x_2 – średnia cena w dniu T_0 ,
- x_3 – średnia cena w tygodniu T_{0-1} (ostatnim tygodniu poprzedzającym dzień T_0 ,
- x_4 – średnia cena w tygodniu T_{0-2} (przedostatnim tygodniu),
- x_5 – średnia cena w tygodniu T_{0-3} (tygodniu poprzedzającym przedostatni tydzień),
- x_6 – średnia cena w okresie T_{0-R-4} (4 tygodniach poprzedzających dany dzień w roku poprzednim),
- x_7 – średnia cena w okresie T_{0-R+4} (4 tygodniach następujących po danym dniu w roku poprzednim),
- y_1 – średnia cena w tygodniu T_{0+1} (horyzont prognozy 1 tydzień),
- y_2 – średnia cena w tygodniu T_{0+2} (horyzont prognozy 2 tygodnie),
- y_3 – średnia cena w tygodniu T_{0+3} (horyzont prognozy 3 tygodnie),
- y_4 – średnia cena w tygodniu T_{0+4} (horyzont prognozy 4 tygodnie).

Ostatecznie uzyskano 1148 wzorców przeznaczonych do uczenia, weryfikacji i testowania sieci neuronowych (dane od 29.09.2001 do 19.11.2004). Ostatnie 287 wzorców (od 7.02.2004) przeznaczono do testowania SSN, natomiast pozostałe podzielono losowo na zbiory uczący (574 wzorce) i walidacyjny (287 wzorców).

Do tworzenia modeli neuronowych użyto programu Statistica Sieci Neuronowe. Dla każdego z czterech modeli prognostycznych testowano 100 różnych sieci zachowując po 10 najlepszych sieci.

W celu określenia modeli pozwalających uzyskiwać najdokładniejsze prognozy obliczono miernik błędu względnego prognozy (mBw):

$$mBw = |\text{średni Bw}| + \text{odchylenie standardowe Bw}$$

gdzie: Bw – błąd względny prognozy $Bw = \frac{Y_{progn} - Y_{rzecz}}{Y_{rzecz}} \cdot 100\%$

Ostatecznego wyboru sieci neuronowej dla poszczególnych modeli prognostycznych (horyzontów czasowych prognozy) dokonano na podstawie wartości miernika mBw dla danych testowych, które nie były wykorzystywane w procesie uczenia.

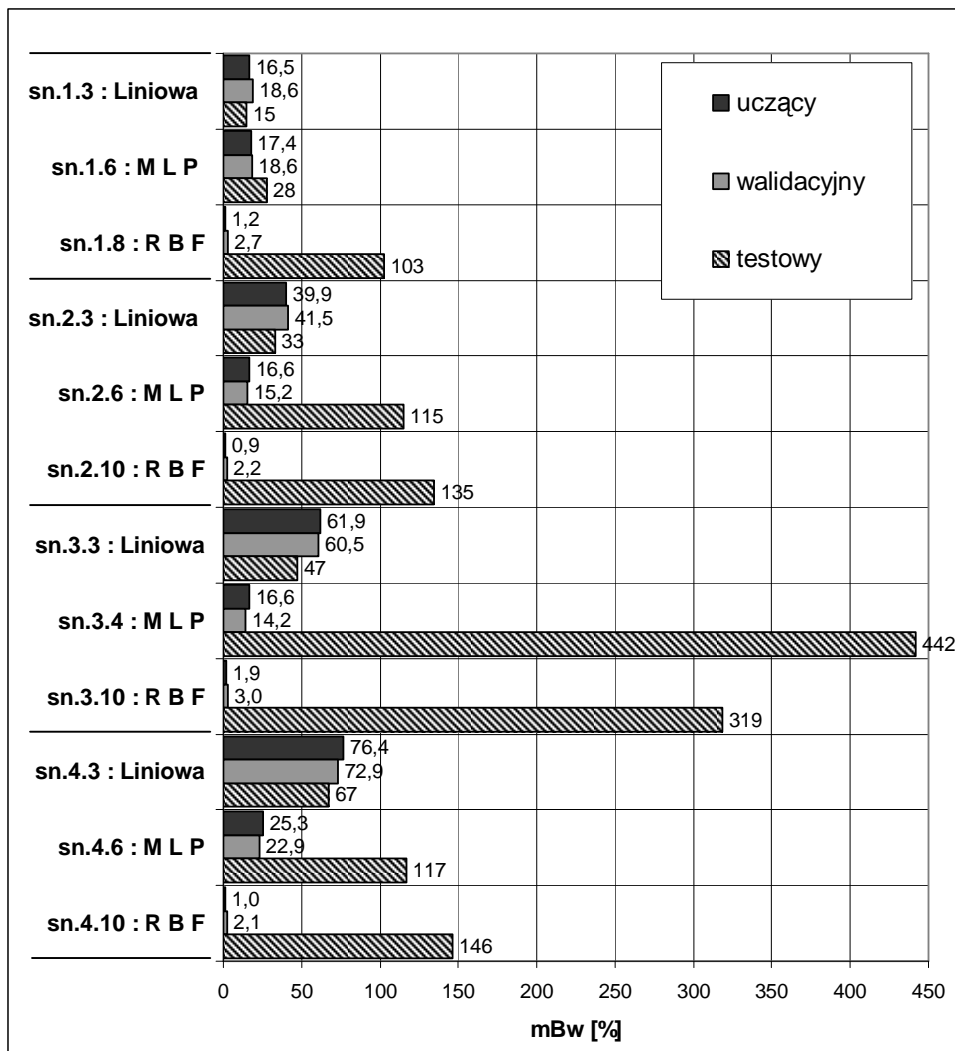
Wyniki badań

Tabela 1. Wartości miernika mBw [%] dla poszczególnych modeli neuronowych (dane walidacyjne)

Table 1. The value of the mBw [%] measurement for the individual neuron models (validating data)

| horyzont prognozy | 1 tydzień | | 2 tygodnie | | 3 tygodnie | | 4 tygodnie | |
|-------------------|-----------|------|------------|------|------------|------|------------|------|
| typ sieci | model 1 | | model 2 | | model 3 | | model 4 | |
| liniowa | sn.1.1 | 23,9 | sn.2.1 | 59,5 | sn.3.1 | 89,8 | sn.4.1 | 86,6 |
| | sn.1.2 | 22,8 | sn.2.2 | 58,2 | sn.3.2 | 88,9 | sn.4.2 | 81,1 |
| | sn.1.3 | 18,6 | sn.2.3 | 41,5 | sn.3.3 | 60,5 | sn.4.3 | 72,9 |
| MLP | sn.1.4 | 20,0 | sn.2.4 | 21,6 | sn.3.4 | 14,2 | sn.4.4 | 41,9 |
| | sn.1.5 | 19,3 | sn.2.5 | 18,1 | sn.3.5 | 17,4 | sn.4.5 | 32,3 |
| | sn.1.6 | 18,6 | sn.2.6 | 15,2 | sn.3.6 | 22,3 | sn.4.6 | 22,9 |
| RBF | sn.1.7 | 4,6 | sn.2.7 | 2,2 | sn.3.7 | 3,8 | sn.4.7 | 5,0 |
| | sn.1.8 | 2,7 | sn.2.8 | 2,3 | sn.3.8 | 4,3 | sn.4.8 | 4,5 |
| | sn.1.9 | 4,5 | sn.2.9 | 2,9 | sn.3.9 | 3,8 | sn.4.9 | 2,2 |
| | sn.1.10 | 3,6 | sn.2.10 | 2,2 | sn.3.10 | 3,0 | sn.4.10 | 2,1 |

W tabeli 1 zestawione zostały wartości miernika mBw dla danych walidacyjnych. Na tej podstawie wybrano najlepsze modele prognostyczne spośród poszczególnych typów sieci neuronowych (rys. 1). Analizując wartości miernika błędu prognozy dla poszczególnych sieci neuronowych (dane uczące, walidacyjne i testowe), można zauważyć, że wielowarstwowe perceptrony (MLP) i sieci o radialnych funkcjach bazowych (RBF) znacznie lepiej opisują charakter zmian cen (mniejsze wartości mBw dla danych uczących i walidacyjnych) niż sieci liniowe (rys. 1). Natomiast wartości mBw dla danych testujących są wielokrotnie większe – szczególnie sieci RBF nie pozwalają na uzyskanie dokładnych prognoz.

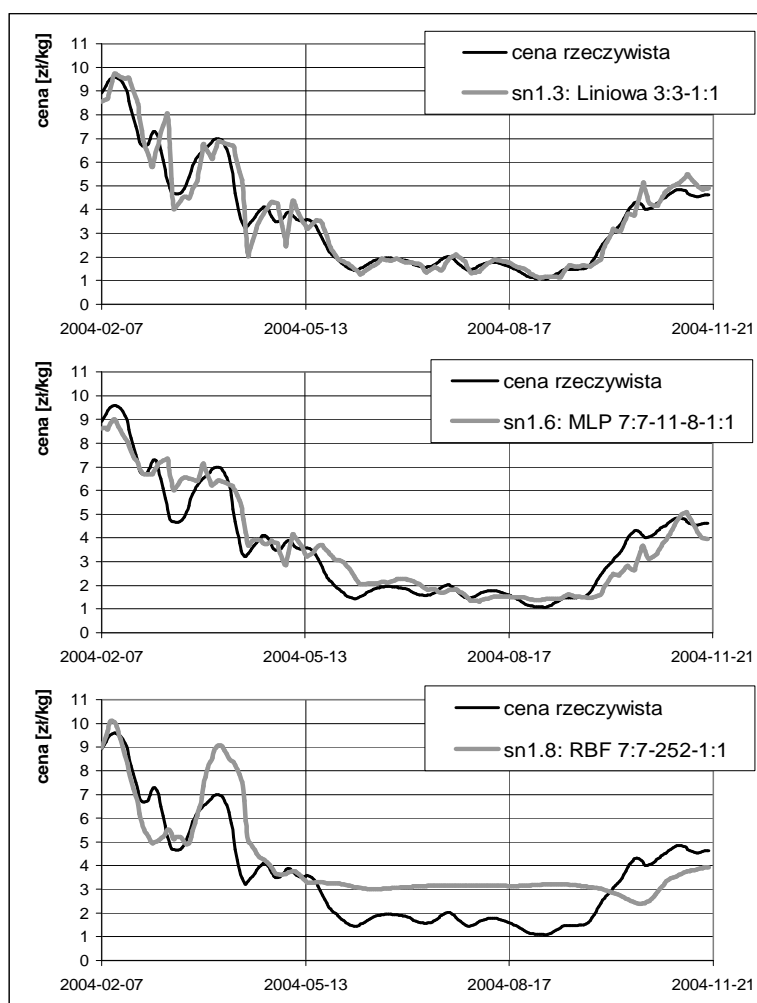


Rys. 1. Wartości miernika mBw dla najlepszych sieci neuronowych
 Fig. 1. The value of the mBw measurement for the best neuron networks.

Wraz z wydłużeniem horyzontu czasowego prognozy rośnie jej błąd – dla danych testujących wartość miernika mBw wynosiła dla sieci:

- liniowych: od 15% (horyzont 1 tydzień) do 67% (horyzont 4 tygodnie),
- MLP: od 28% (horyzont 1 tydzień) do 432% (horyzont 3 tygodnie),
- RBF: od 103% (horyzont 1 tydzień) do 319% (horyzont 3 tygodnie).

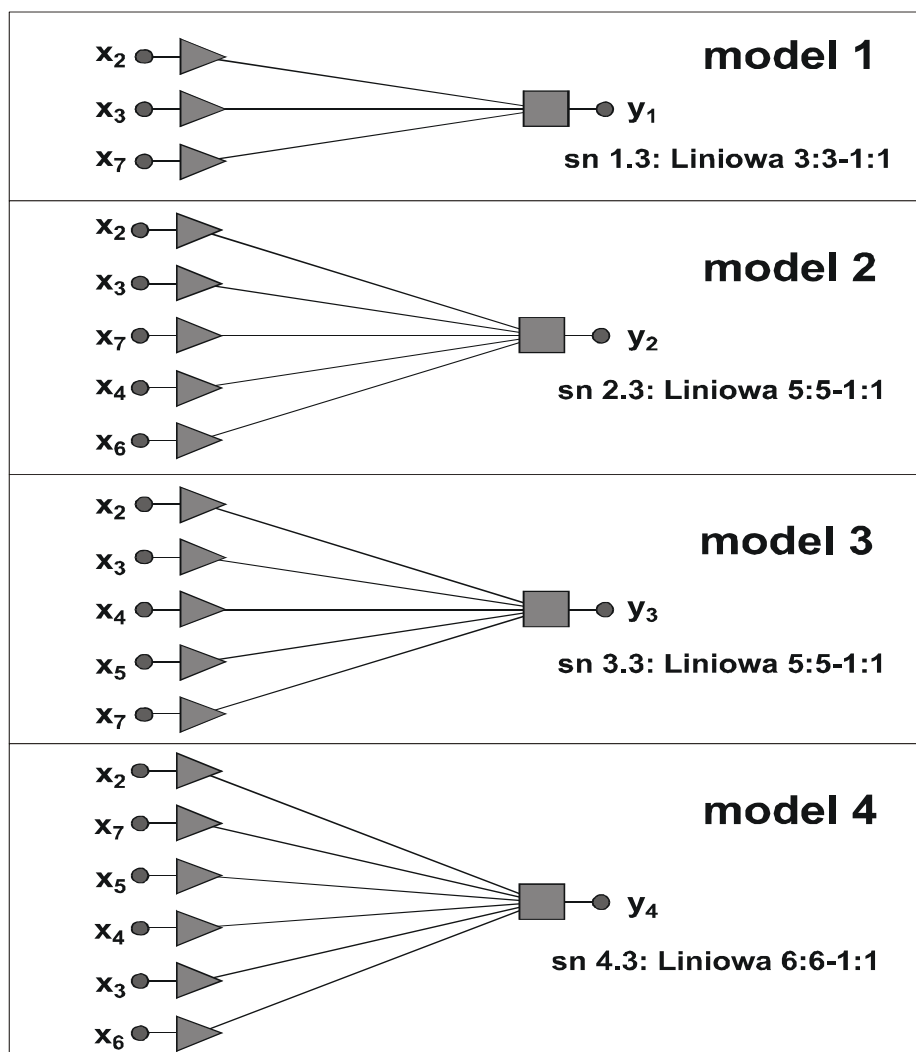
Na rysunku 2 przedstawiono przykładowe wykresy cen rzeczywistych i cen prognozowanych przez najdokładniejsze modele SSN (wybrane na podstawie wartości mBw dla danych walidacyjnych) dla poszczególnych typów sieci (Liniowe, MLP i RBF). Prognozy wykonano dla danych testujących. Horyzont czasowy prognozy wynosił 1 tydzień (model 1). Stwierdzono, że najdokładniej cenę prognozuje sieć liniowa, a najmniej dokładnie sieć typu RBF.



Rys. 2. Porównanie ceny rzeczywistej i prognozowanej przy użyciu SSN dla modelu 1 (horyzont czasowy prognozy 1 tydzień)

Fig. 2. The comparison of the actual price and the price predicted with the use of ANN for the model 1 (1 week time horizon of the forecast)

Podobnie dla pozostałych modeli prognostycznych (modele 2, 3 i 4), najlepsze prognozy można uzyskać stosując sieci liniowe. Sieci typu MLP i RBF, pomimo dokładniejszego dopasowania dla danych uczących i walidacyjnych, wykazują znacznie większe błędy dla danych testujących. Z tego powodu jako modele prognostyczne wybrano sieci liniowe. Charakterystykę najlepszych sieci neuronowych tego typu, dla poszczególnych modeli prognostycznych, pokazano na rysunku 3.



Rys. 3. Charakterystyka najlepszych modeli prognostycznych
 Fig. 3. Characteristics of the best forecasting models

Wnioski

1. Modele neuronowe umożliwiają prognozowanie cen produktów na giełdach rolno – spożywczych z zadowalającą dokładnością.
2. Porównanie różnych typów sztucznych sieci neuronowych pozwala wysnuć wniosek, że najdokładniejsze prognozy uzyskano przy wykorzystaniu liniowych sieci neuronowych.
3. Sieci typu RBF dostarczały najmniej dokładnych prognoz dla danych testujących, pomimo że najdokładniej opisywały zmiany cen dla danych uczących i walidacyjnych. Prawdopodobnie jest to wynikiem zbyt rozbudowanej architektury tych sieci, co spowodowało utratę zdolności do uogólniania.

Bibliografia

Francik S. 2001. Możliwości zwiększenia dokładności prognoz technicznych uzyskiwanych z użyciem sztucznych sieci neuronowych. Inż. Rolnicza, 11(31), 69-75. Warszawa.

Francik S. 2003. Możliwości wykorzystania SSN do prognozowania sprzedaży maszyn rolniczych w warunkach rynkowych. Inż. Rolnicza 12(54), 97-105.

Francik S., Ślipek Z. 2000. Dokładność prognozy technicznej w zależności od architektury SSN. Pr.PIMR., 2, s.64-65.

FORECASTING A HOTHOUSE CUCUMBER PRICE WITH THE USE OF NEURON NETWORKS

Summary

Models using neuron networks to forecast hothouse cucumber prices have been developed in this research, with four forecast horizons. The accuracy of forecasts obtained with the use of various types of neuron networks (linear, multilayer perceptrons and radial base function networks) have been compared. The linear networks have been selected as the best models as they have generated the most accurate forecasts.

Key words: artificial neuron networks, forecasting.