

PORÓWNANIE ALGORYTMÓW UCZENIA SIECI NEURONOWEJ JEDNOKIERUNKOWEJ, Z CZASOWYM OPÓŹNIENIEM, WYKORZYSTANEJ DO PREDYKCJI WARTOŚCI TEMPERATURY POWIETRZA ATMOSFERYCZNEGO

Streszczenie

Przedstawiono wyniki badań wpływu wyboru algorytmu uczenia sieci neuronowej jednokierunkowej, z czasowym opóźnieniem, o topologii perceptronu wielowarstwowego, wykorzystującej w procesie uczenia algorytm wstecznej propagacji błędów, na wyniki predykcji wartości temperatury powietrza atmosferycznego. Stwierdzono, że algorytm uczenia, *br* – *regularyzacja Bayesa*, okazał się jednym z najlepszych pod względem wszystkich analizowanych parametrów oceny przewidywanych wartości temperatur.

Słowa kluczowe: predykcja temperatury, sieci neuronowe, algorytmy uczenia sieci, MATLAB

Wykaz oznaczeń

- n - liczba punktów prognozy czasowej
- err - lokalny bezwzględny błąd predykcji temperatury, °C
- t - czas symulacji, s
- T_e - wartość empiryczna temperatury, °C
- T_p - wartość prognozowana temperatury, °C
- ψ - średni błąd względny prognozy, %

Wprowadzenie

W wielu inżynierskich i agrotechnicznych zastosowaniach ważnym zagadnieniem jest przewidywanie parametrów powietrza atmosferycznego. Predykcja tych wielkości może być wykorzystana np. w procesach przechowywania produktów rolniczych, podejmowania decyzji dotyczących wentylacji budynków inwentarskich. Zależności pomiędzy poszczególnymi wielkościami

opisującymi stan powietrza atmosferycznego są często trudno do sformalizowania w postaci użytecznego modelu matematycznego, dlatego też często w zagadnieniach predykcji wykorzystuje się metody sztucznej inteligencji. Jedną z takich metod są sztuczne sieci neuronowe. Mitall i Zhang [2003] wykorzystali sztuczną sieć neuronową do predykcji w czasie rzeczywistym parametrów psychrometrycznych powietrza atmosferycznego, uzyskując <5% błąd względny dla predykcji temperatury. Shi [2001] wykorzystywał sieci neuronowe do przetwarzania informacji związanych z profilami temperatur atmosfery ziemskiej. Raju [2001] z powodzeniem zastosował algorytmy sieci neuronowych do opisu parametrów pogodowych m.in., temperatury. Ferreira i in. [2002] do przewidywania zmian temperatury w cieplarni. Bodri i Jermak [2003] do przewidywania zmian temperatury gruntu. Często zadanie predykcji rozwiązywane jest na podstawie analizy jednowymiarowego szeregu czasowego, czyli obliczenie wartości prognozowanych (horyzont prognozy) opiera się na podstawie pewnej liczby wartości poprzedzających (rząd prognozy) [Frank i in. 2001]. Między innymi Trajer [2001] wykorzystał tę zasadę do predykcji temperatury otoczenia przy rozważaniach związanych z modelowaniem procesu przechowywania warzyw.

Celem pracy było porównanie algorytmów uczenia jednokierunkowej sieci neuronowej z czasowym opóźnieniem, o topologii perceptronu wielowarstwowego, wykorzystanej do zagadnienia predykcji temperatury powietrza atmosferycznego.

Metodyka badań

W badaniach wykorzystano pomiary temperatury powietrza atmosferycznego wykonane w ciągu 100 dni 1988 r. w Stacji Meteorologicznej w Olsztynie. Rejestracja pomiarów dokonywana była co 3 godziny. Ocenie porównawczej poddano efekty predykcji sieci neuronowej jednokierunkowej, o topologii perceptronu wielowarstwowego z czasowym opóźnieniem wynoszącym 8 pomiarów (rząd prognozy). Horyzont prognozy wynosił 48 godzin, tj. 16 pomiarów.

W symulacjach wykorzystano 12 algorytmów uczenia sieci opartych na zasadzie wstecznej propagacji błędu [Demuth, Beale 2001; Świąć, Bilski 2000]:

- *gd* – algorytm najmniejszego spadku, wagi i bias jest uaktualniany w kierunku ujemnego wektora gradientu funkcji błędu, może być używany w przyrostowej metodzie uczenia,
- *gdm* – zmodyfikowany algorytm *gd*, krok opadania modyfikowany jest o współczynnik *momentum*, pozwalający sieci unikać pułapek utknięcia w minimum lokalnym, może być używany w przyrostowej metodzie uczenia,
- *gdx* – adaptacyjny algorytm szybkości uczenia; może być używany tylko we wsadowej metodzie uczenia,
- *rp* – elastyczny algorytm wstecznej propagacji błędu jest mało czuły na ustawienia parametrów uczenia, dobrze sprawdzający się przy funkcjach sigmoidalnych,
- *cgf (Fletcher-Reeves)* – wykorzystuje algorytmy gradientów sprzężonych zazwyczaj dużo szybsze niż algorytmy ze zmiennym współczynnikiem uczenia, algorytmy gradientów sprzężonych wymagają niewiele więcej pamięci w porównaniu z prostszymi algorytmami, są często dobrym wyborem dla sieci o dużej liczbie wag,
- *cgp (Polak-Ribiere)* – inna wersja algorytmów gradientów sprzężonych, podobnie jak poprzedni wymaga liniowego przeszukiwania w każdej iteracji,
- *cgb (Powell-Beale)* - wersja algorytmów gradientów sprzężonych wykorzystująca niestandardową metodę zmiany kierunku poszukiwania odpowiedniej wartości gradientu,
- *scg* – skalowalny algorytm gradientów sprzężonych nie wykorzystujący metody liniowego przeszukiwania w każdej iteracji, ale kombinowaną metodę algorytmu Levenberga-Marquardta i zbieżności gradientów sprzężonych,
- *bfg* – metoda BFGS quasi-Newtona wymaga zastosowania macierzy hesjana i więcej obliczeń w każdej iteracji w porównaniu z metodą gradientów sprzężonych,
- *oss* – jednokrokowa metoda sieczna, kompromisowa pomiędzy metodami gradientów sprzężonych i quasi-Newtona,
- *lm* – algorytm Levenberga-Marquardta, wyposażony w funkcję redukcji pamięci przy dużych zbiorach uczących,

- *br* – regularyzacja bayesowska, zmodyfikowany algorytm Levenberga-Marquardta, stworzony w celu tworzenia sieci o lepszych własnościach uogólniania, posiadający procedury optymalizacji liczby parametrów sieci.

Do badań wykorzystano oprogramowanie Toolbox Neural Networks v.4 pakietu MATLAB (MathWorks Inc., MA, USA), uruchamiane na komputerze PC wyposażonym w procesor Athlon XP 1900+, 1 GB RAM, pracującym pod kontrolą systemu operacyjnego Windows XP Prof. We wszystkich przeprowadzonych symulacjach sieć miała tę samą strukturę, tj.: 2 neurony w warstwie ukrytej (funkcja przejścia - *tan-sigmoid*), 1 neuron w warstwie wyjściowej (funkcja przejścia - *liniowa*) i te same parametry uczenia sieci: współczynnik uczenia = 0,05, współczynnik momentum = 0,2, liczba epok = 1000, minimalna wartość gradientu = $1,0 \times 10^{-10}$, wartość funkcji celu = $1,0 \times 10^{-10}$ [Demuth, Beale 2001]. Dla potrzeb uczenia sieci dane zostały przeskalowane do zakresu [-1,1].

Jako kryteria oceny algorytmów uczenia wybranego modelu neuronowego przyjęto wartości temperatur: minimalną, maksymalną, średnią, odchylenie standardowe z próby oceniając, na ile te wartości różnią się od danych eksperymentalnych. Dodatkowymi parametrami oceny był współczynnik korelacji *r Pearsona* oraz średni błąd względny prognozy [Dittmann cyt. za Trajer 2001]:

$$\psi = \frac{1}{n} \sum_{i=1}^n \frac{|T_{ei} - T_{pi}|}{T_{ei}} \cdot 100\% \quad (1)$$

oraz lokalny błąd bezwzględny predykcji temperatury opisany formułą (2)

$$err(t) = T_e(t) - T_p(t) \quad (2)$$

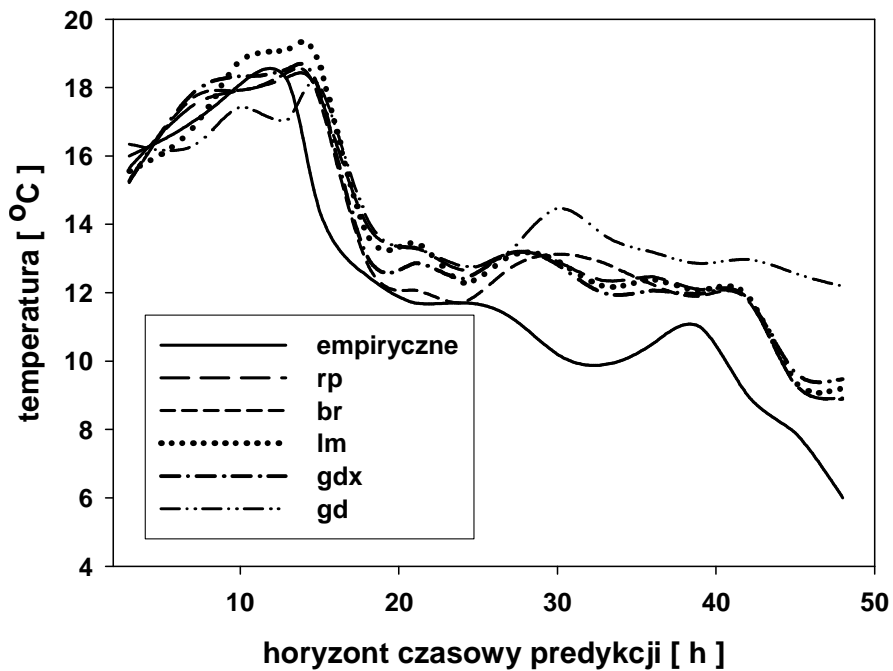
Wyniki badań i ich analiza

Wyniki analizy oceny zastosowania różnych algorytmów uczenia jednokierunkowej sieci neuronowej z czasowym opóźnieniem przedstawione są w tabeli 1.

Tabela 1. Wyniki oceny algorytmów uczenia sieci neuronowej
 Table 1. Results of evaluating the algorithms to education of neural network

	T_{\min} [°C]	T_{\max} [°C]	T_{mean} [°C]	std(T) [°C]	r	ψ [%]	t [s]
empiryczne	6,0	18,2	12,2	3,61			
rp	8,9	18,3	13,8	2,92	0,960	16,1	8,94
cgf	9,2	18,5	13,7	3,01	0,956	15,8	6,5
gdx	9,5	18,5	13,7	2,96	0,951	16,0	8,75
cgb	8,8	18,2	13,4	2,92	0,950	14,0	16,67
scg	9,1	18,8	13,8	3,07	0,945	16,5	13,65
lm	9,2	19,1	13,8	3,1	0,942	16,9	16,96
oss	8,7	18,3	13,4	2,99	0,942	13,8	15,26
cgp	8,9	18,4	13,5	2,99	0,939	14,4	6,19
br	8,9	18,4	13,5	2,99	0,938	14,6	4,68
bfq	8,3	19,3	13,7	3,30	0,930	15,3	16,38
gdm	12,4	17,5	14,4	1,92	0,905	26,0	8,71
gd	12,2	18,1	14,4	1,95	0,875	26,5	8,6

Na podstawie analizy wartości odchylenia standardowego temperatury (dla całego horyzontu prognozy) można stwierdzić, że wszystkie zastosowane algorytmy uczenia „spłaszczają” przewidywane wartości temperatur (wartości odchyłeń standardowych uzyskanych dla 12 algorytmów uczenia są mniejsze niż wartość odchylenia standardowego dla danych empirycznych). Natomiast na podstawie analizy wartości temperatur (minimalnej, maksymalnej i średniej) oraz przebiegów przedstawionych na rysunku 1 można stwierdzić, że wszystkie zastosowane algorytmy uczenia zawyżają przewidywane wartości temperatur (wartości minimalna i średnia temperatury dla danych doświadczalnych jest najmniejsza).

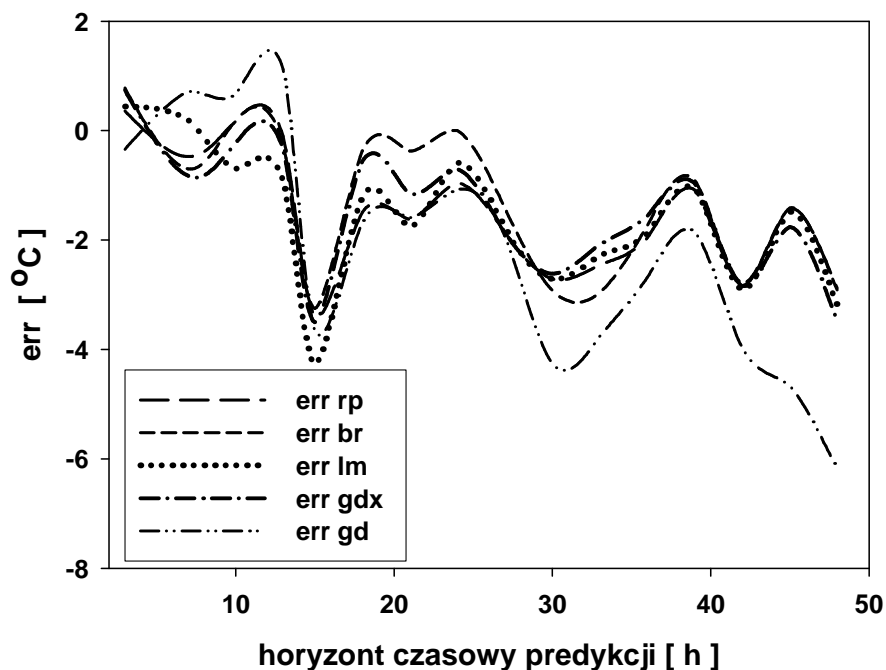


Rys. 1. Doświadczalne i wybrane, przewidywane przebiegi zmian temperatury powietrza atmosferycznego dla sieci neuronowej wykorzystującej różne algorytmy uczenia

Fig. 1. Experimental and selected, predicted courses of atmospherical air temperature changes for neural network using various education algorithms

Dla wszystkich analizowanych przebiegów współczynnik korelacji r Pearsona pokazuje istotną zależność (przy poziomie istotności $p < 0,01$) pomiędzy wartościami przewidywanymi a empirycznymi i wszystkie wartości r z wyjątkiem najgorszego algorytmu gd są $> 0,900$. Wartości średniego względnego błędu prognozy dla 10 algorytmów uczenia są zbliżone do siebie (mieszczą się w przedziale $13,8 \div 16,9\%$) z wyjątkiem dwóch najgorszych gd i gdm , dla których wartość tego współczynnika wynosi ok. 26%. Czas uczenia sieci t jest to czas osiągnięcia przez procedurę uczącą jednego z zadanych parametrów uczenia sieci neuronowej. Najkrótszy czas (4,68 s) odnotowano dla procedury br . Jest on ok. 2÷3 razy krótszy od innych procedur z porównywalnymi wartościami r i ψ .

Na podstawie analizy prezentowanych na rysunku 2 wybranych przebiegów lokalnego błędu bezwzględnego predykcji temperatury powietrza atmosferycznego można stwierdzić ogólny trend wzrostu tego wskaźnika wraz ze wzrostem horyzontu prognozy oraz że zastosowanie algorytmu *br* prowadzi do zastosowania dobrych wyników prognozy szczególnie dla pierwszych 24 godzin. Można przypuszczać, że dysponując kilkuletnimi wynikami pomiarów dokonywanych co 1 h można byłoby uzyskać dużo większą dokładność predykcji.



Rys. 2. Wybrane przebiegi zmian lokalnego błędu bezwzględnego predykcji temperatury dla sieci neuronowej wykorzystującej różne algorytmy uczenia

Fig. 2. Selected change courses of local absolute error in temperature prediction for neural network using various education algorithms

Wnioski

1. Jednokierunkowa sieć neuronowa, o topologii perceptronu wielowarstwowego z czasowym opóźnieniem, wykorzystująca algorytm wstecznej propagacji błędów może być wykorzystana do predykcji temperatury powietrza atmosferycznego.
2. Algorytm uczenia, *br* – regularyzacja Bayesa, okazał się jednym z najlepszych pod względem wszystkich analizowanych parametrów oceny.

Bibliografia

Bodri L., Čermak V. 2003. Prediction of surface air temperatures by neural network, example based on three-year temperature monitoring at Sporilov Station. *Studia Geophysica et Geodaetica*, 47(1): 173-184

Demuth, H., Beale, M. 2001. *Neural Network Toolbox For Use with MATLAB*. The MathWorks, Inc.

Ferreira P.M., Faria E.A., Ruano A.E. 2002. Neural network models in greenhouse air temperature prediction. *Neurocomputing*, 43: 51–75

Frank R.J., Davey N., Hunt S.P. 2001. Time Series Prediction and Neural Networks. *Journal of Intelligent and Robotic Systems*, 31: 91–103

Mittal G.S., Zhang J. 2003. Artificial Neural Network-based Psychrometric Predictor. *Biosystems Engineering*, 85 (3): 283–289, doi:10.1016/S1537-5110(03)00071-0

Raju K.G. 2001. Prediction of Soil Temperature by Rusing Artificial Neural Networks Algorithms. *Nonlinear Analysis*, 47: 1737-1748

Shi L. 2001. Retrieval of Atmospheric Temperature Profiles from AMSU-A Measurement Using a Neural Network Approach. *Journal of Atmospheric and Oceanic Technology*, 18: 340-347

Świąć A., Bilski J. 2000. Metoda wstecznej propagacji błędów i jej modyfikacje. *Biocybernetyka i Inżynieria Biomedyczna*, Tom 6. Sieci Neuronowe. Akademicka Oficyna Wydawnicza EXIT, Warszawa

Trajer J. 2001. Modelowanie procesu przechowywania warzyw w wybranych jego aspektach. Rozprawa habilitacyjna. Szkoła Główna Gospodarstwa Wiejskiego, Warszawa

**COMPARISON OF ALGORITHMS TO EDUCATION
OF UNIDIRECTIONAL NEURAL NETWORK, WITH TIME-
LAG, USED TO PREDICTING VALUES OF
ATMOSPHERICAL
AIR TEMPERATURE**

Summary

Paper presented the study on the effect of selecting an algorithm to education of unidirectional neural network with time-lag, of multilayer perceptron topology, and using an algorithm of reversal error propagation, on results of predicting values of atmospherical air temperature. It was stated that the education algorithm, br – Bayes' regularization, appeared to be one of the best with respect to all analyzed parameters evaluating predicted temperature values.

Key words: temperature prediction, neural networks, algorithms of network education, MATLAB

Recenzent – Jerzy Weres