

METODYCZNE ASPEKTY TWORZENIA MODELI SSN W ZAGADNIENIACH AGROFIZYCZNYCH

Zbigniew Ślipek, Sławomir Francik, Jarosław Frączek

Katedra Podstaw Budowy Maszyn, Akademia Rolnicza, ul. Balicka 104, 30-149 Kraków
e-mail: sfrancik@op.pl

Streszczenie: W pracy przedstawiono etapy tworzenia modelu przy wykorzystaniu sztucznych sieci neuronowych (SSN). Dokonano analizy postępowania podczas kluczowych faz tworzenia modelu na podstawie wyników badań własnych. Sformułowano zalecenia pozwalające na zapewnienie maksymalnej dokładności działania modelu neuronowego. Podczas tworzenia modelu SSN należy przebadać możliwie największą liczbę różnych architektur sieci. Powtarzanie wielokrotnie procesu uczenia zwiększa prawdopodobieństwo znalezienia globalnego minimum błędu działania SSN w hiperprzestrzeni błędów. Wybór najlepszej sieci powinien być przeprowadzony na podstawie wartości błędów dla zbioru testującego, do którego należą wzorce nie biorące udziału w procesie uczenia.

Słowa kluczowe: sieć neuronowa, modelowanie

WSTĘP

Prace naukowo-badawcze, prowadzone w inżynierii rolniczej i agrofizyce, coraz częściej wykorzystują nowoczesne narzędzie modelowania, jakimie dostarczają sztuczne sieci neuronowe (SSN). Szczególnie często stosowane są jednokierunkowe wielowarstwowe SSN [1,2,4,6,10].

Sztuczne sieci neuronowe jako uniwersalny aproksymator mogą być z powodzeniem stosowane do tworzenia modeli empirycznych opisujących różne zjawiska i procesy towarzyszące pozyskiwaniu i przetwarzaniu materiałów roślinnych (biologicznych). W badaniach agrofizycznych występuje duża liczba czynników determinujących przebieg procesów, a wartości cech charakteryzujących materiał biologiczny są bardzo zróżnicowane. Często również brak jest teorii pozwalającej w sposób ścisły opisać przebieg procesu.

Sieci neuronowe pozwalają na opracowanie dobrze działającego modelu bez konieczności tworzenia teorii, a porównanie modeli wykorzystujących SSN z modelami teoretyczno - empirycznymi wykazuje, iż te pierwsze mają większą dokładność. Zastosowanie SSN pozwala ponadto na tworzenie modeli w oparciu o wyniki pomiarów obarczone dużymi błędami losowymi (przetwarzanie zaszumionych wzorców), co w przypadku materiałów biologicznych charakteryzujących się dużą zmiennością cech i trudnościami ich pomiaru, jest jedną z ważniejszych zalet. Dodatkowo sztuczne sieci neuronowe posiadają zdolność uogólniania nabytej wiedzy.

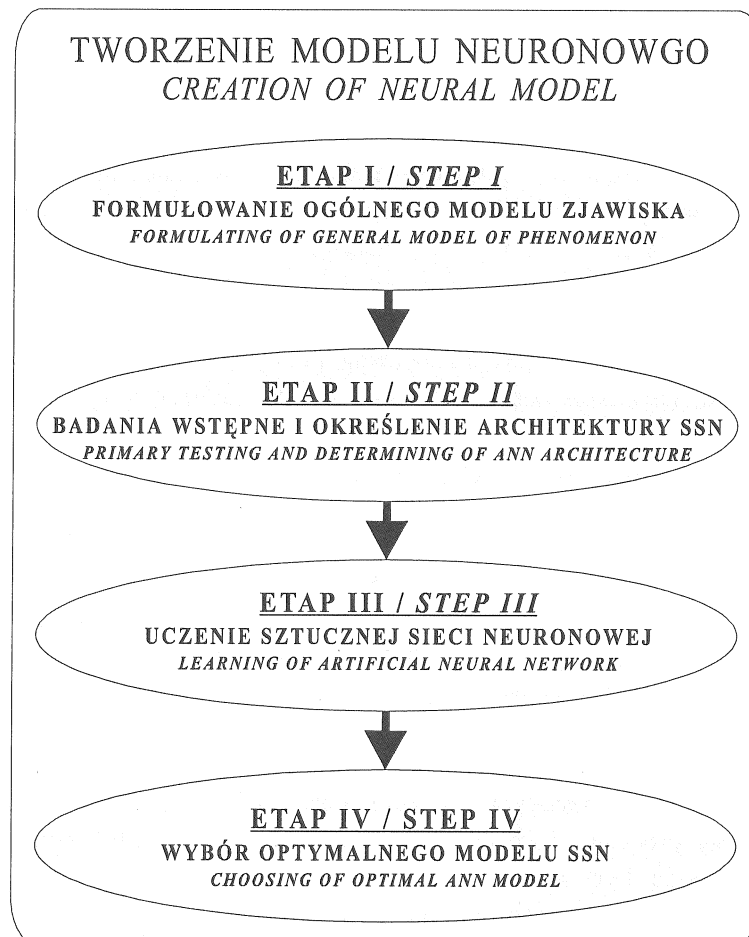
Niestety maksymalne wykorzystanie możliwości SSN jest często nieosiągalne, ze względu na brak jednoznacznych zaleceń sposobu tworzenia modeli. Zarówno dobór rodzaju sieci neuronowej, jak również określenie jej topologii (liczby warstw, liczby neuronów, rodzaju funkcji aktywacji) nie są ujęte w ścisłe ramy. W rezultacie wybór ten jest arbitralny i zależy od twórców modelu. Nie ma zatem pewności, iż otrzymana sieć neuronowa jest najlepszą siecią. Wielokrotne powtarzanie procesu uczenia może doprowadzić do uzyskania lepszego modelu. Jednak poprawa dokładności działania SSN jest najczęściej nieznaczna.

Kolejnym z problemów występujących przy stosowaniu SSN jest występowanie niekorzystnego zjawiska tzw. przeuczenia sieci. Polega ono na utracie przez sieć neuronową zdolności do uogólniania nabytej wiedzy, czego efektem jest wzrost błędu działania sieci dla danych nie należących do wzorców uczących.

Celem pracy było opracowanie metodyki postępowania w tworzeniu modeli zjawisk agrofizycznych przy wykorzystaniu jednokierunkowych sztucznych sieci neuronowych tak, aby zapewnić maksymalną dokładność działania sieci.

ETAPY TWORZENIA MODELU NEURONOWEGO

Analiza wyników badań dotyczących stosowania SSN w technice rolniczej, opublikowanych przez innych naukowców [5,9], jak również wyników badań prowadzonych przez autorów niniejszej pracy [3,8,11], wskazuje, że brak jest ściśle określonych reguł postępowania przy tworzeniu modeli wykorzystujących SSN. Można jednak wyróżnić kilka charakterystycznych etapów, które występują podczas tworzenia modelu (rys. 1).



Rys. 1. Etapy tworzenia modelu wykorzystującego SSN

Fig. 1. Phases of creating a model using ANN

1. **Formułowanie ogólnego modelu (zapis istoty działania modelu).** Przed przystąpieniem do tworzenia modelu neuronowego należy podjąć decyzję dotyczącą liczby i rodzaju zmiennych zależnych (wyjść sieci neuronowej) i zmiennych niezależnych (wejść do sieci).
2. **Badania wstępne i określenie architektury SSN.** Podczas tego etapu dzieli się losowo dane na zbiory uczący i testujący (ewentualnie weryfikujący), oraz określa liczbę warstw i neuronów w warstwach możliwą do zastosowania. Podejmuje się również decyzję o rodzaju funkcji aktywacji neuronów. Można również przeprowadzić normalizację (skalowanie) zmiennych wejściowych i wyjściowych.

3. **Uczenie SSN.** Proces uczenia poprzedza wybór metody uczenia sieci neuronowej, oraz określenie wartości parametrów tego procesu. Podejmuje się również decyzję dotyczącą liczby epok uczenia.
4. **Wybór optymalnego modelu SSN.** Po wcześniejszym przyjęciu kryterium optymalizacji, ustala się optymalną architekturę sieci neuronowej i określa, która z SSN najlepiej modeluje badany proces. Można następnie kontynuować proces uczenia dla wybranej sieci neuronowej, aby zwiększyć dokładność działania modelu.

Wymienione cztery etapy tworzenia modelu mogą być realizowane w rozmaity sposób. Poniżej podano propozycje oraz uwagi praktyczne dotyczące realizacji tych etapów, wynikające z doświadczeń autorów w praktycznym stosowaniu SSN.

FORMUŁOWANIE OGÓLNEGO MODELU ZJAWISKA

Wybór rodzaju zmiennych wejściowych i wyjściowych SSN wynika zazwyczaj ze znajomości teoretycznych modeli opisywanego zjawiska, lub zależności empirycznych, względnie z postawionych hipotez badawczych. Liczba zmiennych wejściowych do SSN może być praktycznie dowolna, należy jednak unikać ich nadmiernego mnożenia, gdyż powoduje to konieczność zwiększenia liczebności zbioru uczącego. Zwiększa to koszty badań empirycznych koniecznych do uzyskania wystarczającej liczby wzorców uczących. Z drugiej strony zwiększenie liczby zmiennych wejściowych może umożliwić lepszy opis modelowanego procesu i spowodować zwiększenie dokładności działania SSN. Przykładem mogą być wyniki badań autorów dotyczące stworzenia modelu zjawiska tarcia zewnętrznego ziarnistych materiałów roślinnych [3]. Porównywano tam trzy modele neuronowe o różnej liczbie zmiennych wejściowych. Najbardziej rozbudowany model uzyskano dla 8 zmiennych wejściowych. Była to sieć neuronowa o 15 neuronach w warstwie pierwszej i drugiej i jednym neuronie w warstwie wyjściowej, dla której średni błąd względny wynosił – 0,22% a odchylenie standardowe tego błędu 22,47% (dla danych testujących). Dla drugiego modelu, o 7 zmiennych wejściowych, średni błąd względny wynosił – 2,10% a odchylenie standardowe 23,41%. Sieć ta miała również dość rozbudowaną architekturę: 11 neuronów w pierwszej warstwie, 13 neuronów w drugiej warstwie i 1 neuron w warstwie wyjściowej. Trzeci z opracowanych modeli miał jedynie 3 zmienne wejściowe, a jego architektura była identyczna jak modelu pierwszego. Średni błąd względny dla trzeciego modelu wyniósł – 0,59%, a odchylenie standardowe tego błędu 24,92%. A zatem zmniejszenie liczby

zmiennych wejściowych z 8 do 7 spowodowało pogorszenie dokładności działania modelu, natomiast dalsze zmniejszenie liczby zmiennych wejściowych do 3 nie zmniejszyło znacząco dokładności działania modelu.

Wybór liczby zmiennych wyjściowych jest zdeterminowany celem modelowania. Jednak w przypadkach, kiedy chcemy uzyskiwać na wyjściu więcej niż jedną wartość, korzystniejsze jest stworzenie kilku modeli neuronowych, każdy o jednym wyjściu, działających niezależnie, niż jednej sieci neuronowej o kilku wyjściach.

BADANIA WSTĘPNE I OKREŚLENIE ARCHITEKTURY SSN

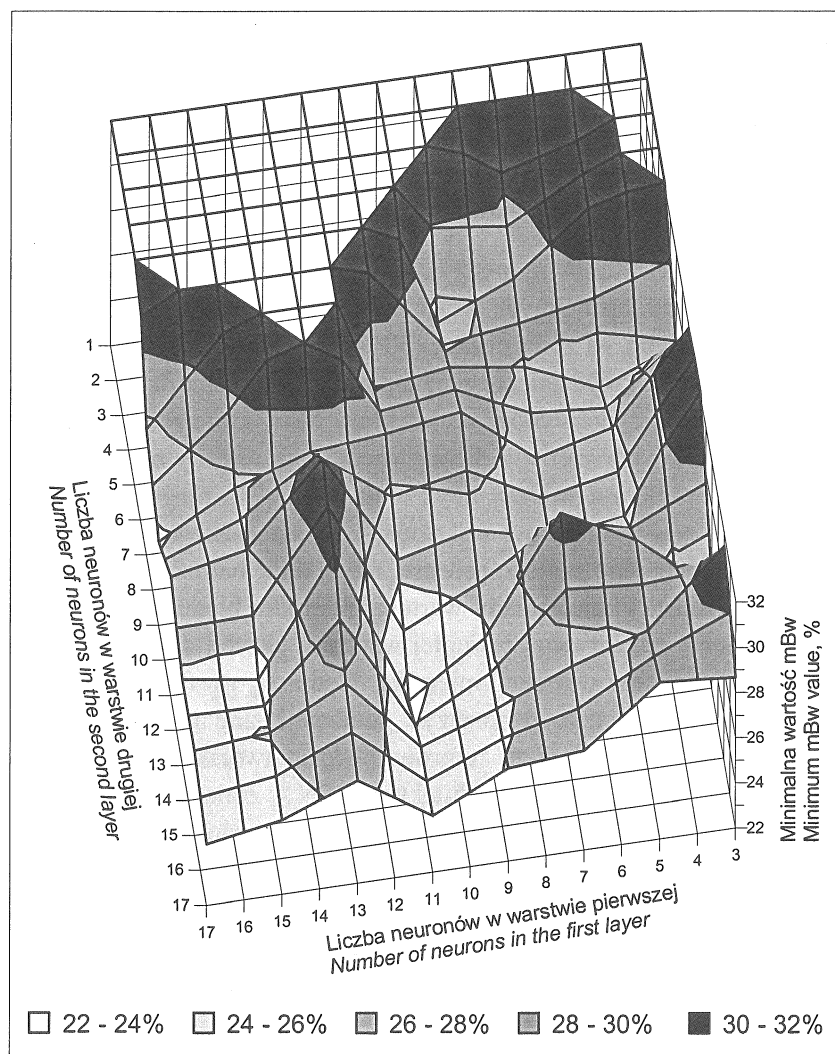
Drugi etap modelowania rozpoczyna się od podziału posiadanych danych empirycznych, wartości zmiennych wejściowych i odpowiadających im wartości zmiennych wyjściowych, na zbiór uczący (50÷70% wzorców uczących) i testujący. Niekiedy wymagane jest wydzielenie ze zbioru testującego zbioru weryfikującego.

Następnie dokonuje się wyboru architektury SSN, która ma stanowić model badanego procesu. Wybór zarówno liczby warstw jak i liczby neuronów w poszczególnych warstwach jest trudny, gdyż brak jest co do niego jednoznacznych zaleceń. Badania własne autorów wskazują, że sieci różniące się tylko 1 neuronem, użyte do opisu tego samego procesu, mogą mieć bardzo różną dokładność. Na rysunku 2 przedstawiono zmiany wartości miernika błędu mBw w zależności od liczby neuronów w pierwszej i drugiej warstwie ukrytej dla danych ze zbioru testującego [2]. Zmiany wartości miernika błędu są gwałtowne i trudne do przewidzenia – nie można wyznaczyć tendencji zmian dokładności sieci neuronowej. Dlatego też, zdaniem autorów, podczas tworzenia modelu powinno się przebadać wszystkie możliwe architektury SSN – zmieniać liczby neuronów w poszczególnych warstwach w jak najszerszych granicach.

Zakres zmian liczby warstw i liczby neuronów jest ograniczony liczbą wzorców uczących. Generalnie im bardziej rozbudowana sieć neuronowa, tym większa jest liczba wag tej sieci i tym liczniejszy musi być zbiór uczący. Liczba wag SSN wynika z liczby zmiennych wejściowych i wyjściowych, liczby warstw, oraz liczby neuronów w poszczególnych warstwach sieci.

W etapie tym można również przeprowadzić skalowanie wartości zmiennych wejściowych i wyjściowych. Skalowanie zmiennych wejściowych powinno być przeprowadzane, gdy wartości poszczególnych zmiennych różnią się znacznie, co może powodować trudności w uczeniu się sieci neuronowej. Natomiast skalowanie zmiennych wyjściowych jest konieczne, gdy zakres wartości wyjściowych generowanych przez sieć neuronową jest ograniczony. Jest to związane z rodzajem funkcji

aktywacji neuronów warstwy wyjściowej. Tak np., często stosowana funkcja sigmoidalna unipolarna zmienia wartość od 0 do 1, a zatem nie można uzyskać na wyjściu sieci wartości większej niż 1.



Rys. 2. Wartości minimalne miernika *mBw* w zależności od architektury sieci neuronowych dla danych testujących

Fig. 2. Minimal values of the *mBw* measure depending on the neural network architecture for the testing data

UCZENIE SZTUCZNEJ SIECI NEURONOWEJ

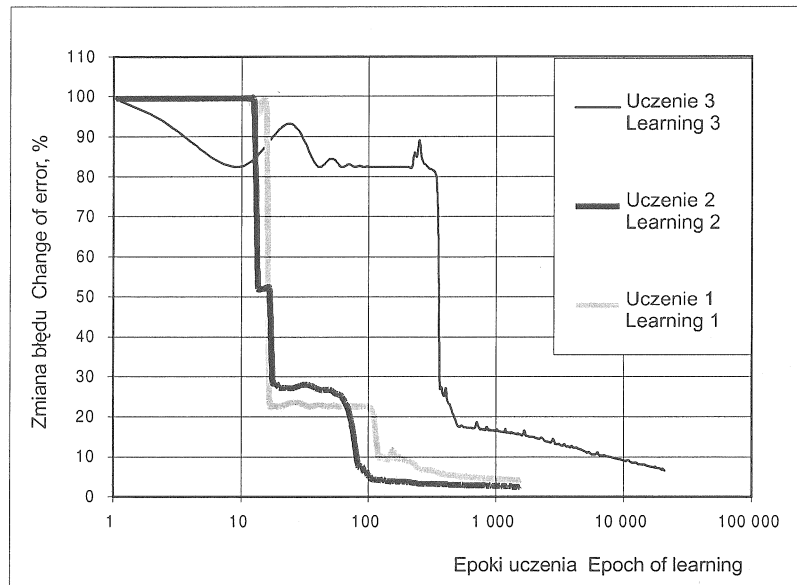
Trzeci etap tworzenia modelu neuronowego rozpoczyna się od wyboru algorytmu uczenia. Dotychczas najczęściej stosowanym algorytmem dla sieci jednokierunkowych była metoda wstecznej propagacji błędów. Ponieważ przeważnie korzysta się z gotowego oprogramowania do tworzenia modeli neuronowych, metoda uczenia jest często narzucona przez twórców tego oprogramowania. Oprócz metody uczenia ustala się również warunki zakończenia procesu uczenia – określa maksymalną liczbę epok uczenia sieci. Niestety nie można apriori stwierdzić, ile epok stanowi wystarczającą liczbę, powyżej której uczenie nie poprawia dokładności działania SSN. Dotychczasowe badania autorów wskazują, że liczba epok uczenia zmienia się w zależności od modelowanego procesu od kilkuset do kilkudziesięciu tysięcy epok. Liczbę tę należy jednak każdorazowo dostosować do przebiegu procesu uczenia. Na rysunku 3 przedstawiono przykłady zmian błędów działania sieci neuronowych podczas ich uczenia. Jako 100% przyjęto błąd w chwili rozpoczęcia procesu uczenia. Wykresy sporządzono dla modeli opisujących:

- uczenie 1 – proces tarcia zewnętrznego ziarnistych materiałów roślinnych – model I o 7 wejściach [3],
- uczenie 2 – proces tarcia zewnętrznego ziarnistych materiałów roślinnych – model II o 8 wejściach [3],
- uczenie 3 – zmiany wielkości powierzchni buraka ćwikłowego podczas suszenia [7].

Prowadząc badania dotyczące wykorzystania SSN można również zauważyć, że nie tylko sieci różniące się liczbą neuronów charakteryzują się różną dokładnością działania, ale również sieci o tej samej architekturze mogą wykazywać znaczne różnice w dokładności działania. Dlatego też autorzy kilkakrotnie powtarzają proces uczenia sieci neuronowej dla tej samej liczby neuronów (architektury). Na rysunku 4 przedstawiono różnice dokładności modelowania, jakie osiągały SSN o identycznej architekturze, po przeprowadzonym wielokrotnie procesie uczenia:

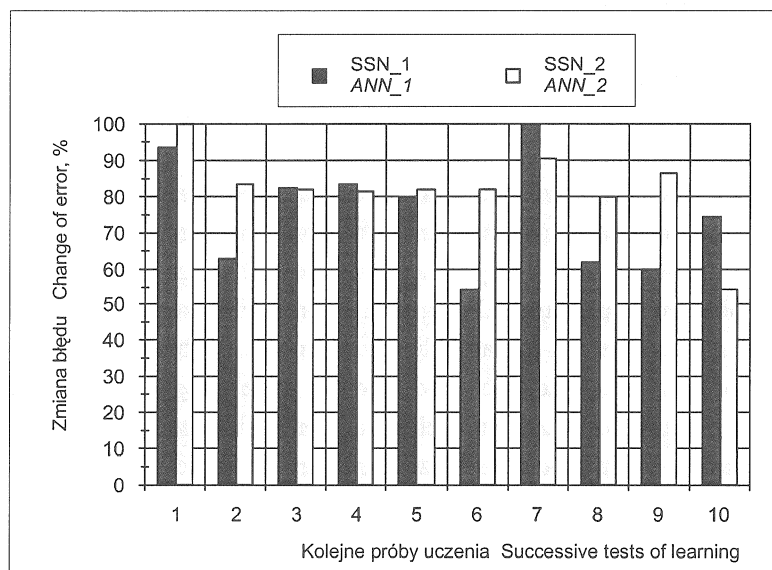
- SSN_1 modelowała proces tarcia zewnętrznego – model I [3],
- SSN_2 opisywała zjawisko uszkodzenia ziarna zbóż [11].

Dla uzyskanych w procesie uczenia sieci neuronowych opisujących badane zjawiska jako 100% przyjęto największy błąd i względem niego obliczono procentowe zmniejszenie błędów dla pozostałych SSN.



Rys. 3. Zmiany wartości błędu podczas uczenia sieci neuronowych

Fig. 3. Changes in the error value in the neural networks learning process



Rys. 4. Porównanie dokładności działania sieci neuronowych uzyskanych podczas kolejnych powtórzeń procesu uczenia

Fig. 4. The comparison of the neuron networks performance accuracy achieved in subsequent repetitions of the learning process

W obu przypadkach najlepsza sieć miała błąd prawie dwukrotnie mniejszy niż najgorsza. Dla pierwszego z przedstawionych procesów (SSN_1), uczenie nr 6 pozwoliło zmniejszyć błąd działania sieci neuronowej do 54%, w porównaniu z siecią uzyskaną w wyniku uczenia nr 7. Dla drugiego z przedstawionych procesów (SSN_2) proporcje te były takie same. Błąd po uczeniu nr 10 stanowił 54% błędu uczenia nr 1.

WYBÓR OPTIMALNEGO MODELU SSN

Czwarty etap procesu tworzenia modelu neuronowego, mający za zadanie wybór najlepszej SSN, musi być poprzedzony przyjęciem kryterium optymalizacyjnego. Kryterium służącym do oceny dokładności działania SSN może być wartość błędu względnego. Ponieważ w ocenie dokładności działania modelu powinno się uwzględnić zarówno miarę tendencji centralnej, jak i miarę rozrzutu, autorzy zaproponowali jako kryterium miernik błędu mBw [8]. Miernik mBw jest obliczany, dla danych testujących, jako suma wartości bezwzględnej średniego błędu względnego oraz odchylenia standardowego tego błędu:

$$mBw = |srBw| + sdBw$$

gdzie: $|srBw|$ – średnia arytmetyczna względnych błędów procentowych,
 $sdBw$ – odchylenie standardowe względnych błędów procentowych.

Wyboru modelu dokonuje się na podstawie wartości miernika mBw dla danych testujących (nie biorących udziału w procesie uczenia). Najlepszą siecią jest sieć o najmniejszej wartości mBw , przy spełnionym warunku wystarczającej liczby wzorców testujących.

PODSUMOWANIE I WNIOSKI

Prezentowany sposób postępowania podczas tworzenia modeli SSN został zweryfikowany pozytywnie podczas badań prowadzonych przez autorów. Umożliwia on uzyskiwanie sieci neuronowych, których działanie jest obciążone znacznie mniejszym błędem niż inne modele tworzone według dotychczas istniejących metod. Można sformułować zalecenia umożliwiające tworzenie modeli neuronowych opisujących badane procesy:

1. Podczas tworzenia modelu SSN należy przebadać możliwie największą liczbę różnych architektur sieci. Należy poddać analizie zarówno sieci dwu jak i trójwarstwowe. Rozmiar sieci jest ograniczony wielkością zbioru uczącego.

2. Powtarzanie wielokrotnie procesu uczenia zwiększa prawdopodobieństwo znalezienia globalnego minimum błędu działania SSN w hiperprzestrzeni błędu. Proces uczenia powinien być prowadzony do momentu, gdy sieć neuronowa przestaje zwiększać dokładność działania.

3. Wybór najlepszej sieci powinien być przeprowadzony na podstawie wartości błędów dla zbioru testującego, do którego należą wzorce nie biorące udziału w procesie uczenia. Pozwala to na uzyskanie modelu posiadającego zdolność do uogólniania nabytej przez sieć wiedzy (uniknięcie zjawiska przeuczenia).

4. Ponieważ proces uczenia jest losowy, nie możemy być pewni, że uzyskany model sieciowy jest optymalny. Może się okazać, że ponowne przeprowadzenie procesu uczenia, dla innych wag początkowych, pozwoli na uzyskanie dokładniej działającej SSN.

PIŚMIENNICTWO

1. **Bruton J.M., McClendon R.W., Hoogenboom G.:** An artificial neural network model of pan evaporation. An ASAE Meeting Presentation, Paper No. 983180, 1998.
2. **Francik S., Frączek J.:** Model development of the external friction of granular vegetable materials on the basis of artificial neural networks. *Int. Agrophysics*, 15, 231-236, 2001.
3. **Francik S., Frączek J.:** Wpływ kształtu pojedynczych nasion na wartość siły tarcia zewnętrznego. *Inż. Rolnicza*, 4 (37), 83-89, 2002.
4. **Jaszuk M., Strobel W.:** Wykorzystanie sztucznej sieci neuronowej przy ocenie siły potrzebnej na rozerwanie strąków łubinu. II Zjazd Naukowy PTA, Referaty i doniesienia, 181-182, Lublin 2000.
5. **Langman J.:** Zastosowanie sztucznych sieci neuronowych w inżynierii rolniczej. *Inż. Rolnicza*, 1(7), 153-158, Warszawa 1999.
6. **Liao K., Paulsen M.R., Reid J.F., Ni B.C., Bonifacio-Maghirang E.P.:** Corn kernel breakeage classification by machine vision using a neural network classifier. *Transactions of the ASAE*, vol. 36(6), 1949-1953, 1993.
7. **Łapczyńska-Kordon B., Francik S.:** Sztuczne sieci neuronowe w modelowaniu zmian wymiarów geometrycznych warzyw podczas konwekcyjnego suszenia. *Prace Komisji Nauk Rolniczych PAU*, nr 2, 83-89, 2001.
8. **Ślipek Z., Francik S.:** Prognozowanie zmian techniczno-eksploatacyjnych parametrów maszyn rolniczych przy użyciu sztucznych sieci neuronowych. *Zesz. Nauk. AR w Krakowie* nr 385, ser. Technika Rolnicza, 17, 29-35, 2001.
9. **Trajer J., Górnicki K.:** Prognozowanie zmian ceny i strat masowych warzyw podczas ich przechowywania z wykorzystaniem sieci neuronowych (w: *Metody komputerowej analizy procesów suszenia i przechowywania warzyw*). Wydawnictwo SGGW, Warszawa 1999.
10. **Yang C.C., Lacroix R., Prasher S.O.:** The use of back-propagation neural networks for the simulation and analyses of time-series data in subsurface drainage systems. *Transactions of the ASAE*, vol. 41(4), 1181-1187, 1998.
11. **Złobecki A., Francik S.:** Defining the damaging process of cereal grains on the basis of artificial neural network. *Int. Agrophysics*, 15, 219-223, 2001.

METHODIC ASPECTS OF CREATING ANN MODELS
IN AGROPHYSICAL RESEARCH

Zbigniew Ślipek, Sławomir Francik, Jarosław Frączek

Department of Machine Design, University of Agriculture
ul. Balicka 104, 30-149 Kraków
e-mail: sfrancik@op.pl

Abstract. The paper presents phases of creating a model with the use of artificial neural networks (ANN). The authors have analysed the key phases of model creation based on own research results. Recommendations that allow to ensure the maximum accuracy of the neural model performance were formulated: During creation of the SSN model it is recommended to test different net architectures as most as possible. Multiple-repeating of learning process increases probability to find a minimum global error of working of SSN in error hyperspace. Choosing the best net shall be carried out basing on values of mistakes for testing set to which belongs standards not taken into account during learning process.

Key words: artificial neural network, modelling

