



**INSTYTUT BADAŃ SYSTEMOWYCH
POLSKIEJ AKADEMII NAUK**

**ANALIZA SYSTEMOWA W FINANSACH
I ZARZĄDZANIU**

Wybrane problemy
Tom 11

Pod redakcją
Jerzego HOŁUBCA

Warszawa 2009



**INSTYTUT BADAŃ SYSTEMOWYCH
POLSKIEJ AKADEMII NAUK**

**ANALIZA SYSTEMOWA W FINANSACH
I ZARZĄDZANIU**

Wybrane problemy
Tom 11

Pod redakcją
Jerzego HOŁUBCA

Warszawa 2009

Wykaz opiniodawców artykułów zamieszczonych
w niniejszym tomie:

prof. dr hab. inż. Jerzy HOŁUBIEC

dr inż. Lech KRUŚ

doc. dr hab. inż. Wiesław KRAJEWSKI

doc. dr hab. Jacek MALINOWSKI

dr inż. Edward MICHALEWSKI

prof. dr Adam SKOREK

dr hab. Ryszard SMARZEWSKI

prof. dr hab. inż. Andrzej STRASZAK

dr Dominik ŚLĘZAK

prof. dr hab. inż. Stanisław WALUKIEWICZ

doc. dr hab. Sławomir ZADROŻNY

© Instytut Badań Systemowych PAN
Warszawa 2009

ISBN 9788389475220

Druk: Zakład Poligraficzny Jerzy Kosiński, Warszawa

ZASTOSOWANIE SZTUCZNYCH SIECI NEURONOWYCH DO POMIARU KAPITAŁU LUDZKIEGO

Anet Anna Wiktorzak

Studia Doktoranckie IBS PAN

Artykuł jest wstępem do badań autorki nad możliwością wykorzystania sztucznych sieci neuronowych w diagnostyce edukacyjnej. Po raz pierwszy zaproponowano pomiar kapitału ludzkiego ucznia szkoły ponadgimnazjalnej jedną z metod sztucznej inteligencji. Zastosowanie modelu sztucznej sieci neuronowej może stanowić przełom w tego typu analizach i badaniach. W artykule znajdziemy podstawy teoretyczne matematycznego modelu neuronowego oraz przykład zbudowania całej sieci.

Słowa kluczowe: sztuczna sieć neuronowa (SSN), uczenie sieci, algorytm uczenia sieci, algorytm wstecznej propagacji błędów, kapitał ludzki (KL), kapitał społeczny (KS).

Wstęp

Jednym z najważniejszych elementów zarządzania szkołą jest dbałość o prawidłowy przebieg procesu edukacyjnego/dydaktycznego. Działania te kształtują, budują i rozwijają wiedzę, umiejętności, osobowość itp. danego ucznia, co będziemy nazywali jego kapitałem ludzkim. O jakości tych procesów decyduje również otoczenie szkoły, ale przede wszystkim nauczyciele: doświadczenie, wiedza, talent pedagogiczny, umiejętności itp., każdego nauczyciela traktowanego jako pojedyncza istota ludzka, co będziemy nazywali kapitałem ludzkim danego nauczyciela oraz ich zdolności do współpracy, wzajemne zaufanie itp., które łącznie tworzą kapitał społeczny danej szkoły. To właśnie szkoła średnia/ponadgimnazjalna jest jednym z najważniejszych etapów budowania kapitału ludzkiego ucznia i przyszłego świadomego obywatela. Zaproponowany w artykule nowy model do szacowania poziomów kapitału ludzkiego pozwala przede wszystkim, na identyfikację i pomiar tych niematerialnych zasobów. Analiza systemowa, metody statystyczne oraz sztuczne sieci neuronowe są dobrymi narzędziami do szacowania tych wielkości, dobrym punktem startu do badań nad ich pomiarami.

1. Sztuczne sieci neuronowe

Sztuczna sieć neuronowa (SSN) to ogólna nazwa struktur matematycznych oraz ich programowych lub sprzętowych modeli, realizujących obliczenia lub przetwarzanie sygnałów poprzez zbiór elementów, zwanych sztucznymi neuronami, które wykonują pewną podstawową operację z sygnałami/informacjami wejściowymi. Jednym z głównych czynników mających wpływ na duże zainteresowanie SSN, jest możliwość ich stosowania w bardzo wielu dziedzinach życia do rozwiązywania problemów, gdzie użycie innych metod jest trudne lub wręcz dzisiaj niemożliwe. Znakomicie sprawdzają się one w problemach klasyfikacji, predykcji, związanych ze sterowaniem, analizą danych, itp.

Przygotowanie sieci do pracy jest dwuetapowe. Pierwszym z nich jest proces uczenia, w którym sieć uczy się na podstawie danych empirycznych jak reagować na zadany bodziec/sygnał. Gdy sieć zostanie nauczona/wytrenowana można przejść do drugiego etapu – procesu egzaminowania, podając na jej wejścia dowolne sygnały, które sieć powinna rozpoznać lub zaklasyfikować.

2. Historia rozwoju sztucznych sieci neuronowych

Prace nad poznaniem procesów zachodzących w mózgu, które mają wpływ na sposób jego działania, były prowadzone od bardzo dawna przez wielu badaczy z całego świata. Przemyslenia te były też podwaliną prac nad SSN. Za początek sztucznych sieci neuronowych przyjmuje się przedstawienie matematycznego modelu neuronu przez Warrena McCulloch'a i Waltera Pitts'a w 1943 roku, które stało się inspiracją dla wielu późniejszych badaczy. W 1949 roku Donald Hebb odkrył, że informacja może być przechowywana jako wartość wag pomiędzy połączeniami poszczególnych neuronów i zaproponował pierwszy algorytm uczenia takiej sieci. Polegał on na zmianach wag połączeń - później został on nazwany regułą Hebba. Pierwszą znaną i opisaną, działającą siecią neuronową był perceptron, opracowany przez Franka Rosenblatt'a w 1957 roku w Cornell Aeronautical Laboratory. W 1960 roku powstała struktura Adaline, na którą składał się pojedynczy element liniowy. Jej twórcą był Bernard Widrow ze Stanford University. Element Adaline po rozbudowaniu o kolejne elementy połączone ze sobą tworzyła sieć Madaline¹.

Początkowe gwałtowne zainteresowanie i postęp w tej dziedzinie został ostudzony w latach 70-tych przez publikację pracy Marvin'a Minsky'ego

¹ Tadeusiewicz R.: *Sieci neuronowe*, Akademicka Oficyna Wydawnicza, Warszawa, 1993, str. 8.

i Seymoura Paperta w 1969 r., która dowodziła, że sieć składająca się z jednej warstwy liniowej ma bardzo ograniczone możliwości. Pomimo tego powstawało wiele nowych pomysłów i modeli sieci. Można tu wymienić sieć Cerebellatron Davida Mara (sterowanie robotem) oraz działająca jak pamięć asocjacyjna Brain in the Box Jamesa Andersona z 1977 r.

Zastój w badaniach nad SSN został przełamany po publikacji algorytmu uczenia wielowarstwowej nieliniowej sieci neuronowej. Metoda wstecznej propagacji błędów została po raz pierwszy opisana przez Paula Werbosa w 1974 roku jednak jej popularyzacja nastąpiła dopiero w 1986 roku za sprawą Davida Rumelharta. To umożliwiło konstruowanie wielowarstwowych nieliniowych sieci neuronowych i ich skuteczne uczenie. Spowodowało to powrót zainteresowania sieciami neuronowymi i dalszy szybki ich rozwój.

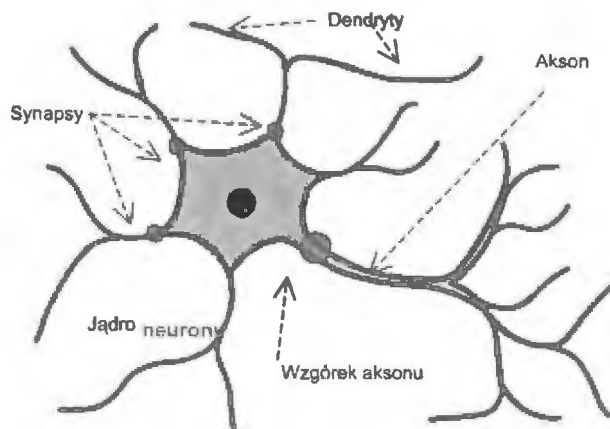
3. Biologiczne inspiracje SSN

Sieci neuronowe są sztucznymi strukturami, których budowa i działanie zostały zaprojektowane w sposób modelujący działanie naturalnego układu nerwowego, w szczególności mózgu. Mózg składa się z komórek nerwowych zwanych neuronami, pomiędzy którymi znajdują się cieniutkie połączenia - tzw. aksony. W powiększeniu fragment mózgu wygląda jak drobna sieć, której węzłami są neurony (patrz rys. 1).

Mózg składa się z około 10 miliardów elementarnych komórek nerwowych nazywanych neuronami. Tworzą one bardzo skomplikowane sieci powiązań między sobą. Neurony w mózgu pracują równolegle. Składają się one z następujących części:

- - Dendrytów - biologiczne neurony mają ich tysiące; dendryty dochodzą do neuronów poprzez synapsy - nośnik pamięci, biochemiczne złącze, modyfikujące sygnały.
- - Jądra - "centrum obliczeniowego" neuronu; to tutaj zachodzą procesy kluczowe dla funkcjonowania neuronu.
- - Aksonu - "wyjście" neuronu, za jego pośrednictwem neuron powiadamia świat zewnętrzny o swojej reakcji na dane wejściowe. Neuron ma tylko jeden akson, rozpoczynający się wzgórkiem aksonu - stąd wysyłany jest sygnał wyjściowy, który wędruje dalej poprzez akson i przechodzi w następne dendryty.

Na podstawie budowy biologicznej powstał model sztucznego neuronu oraz sieci neuronowej. Jest on znacznie uproszczony w stosunku do biologicznego oryginału.



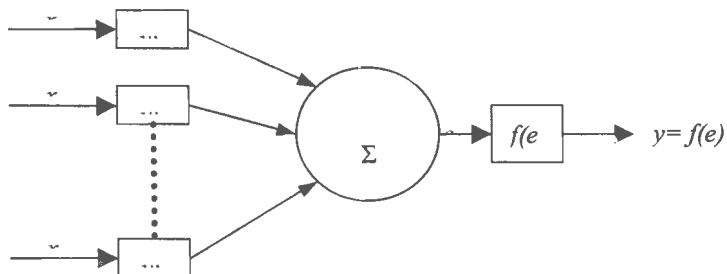
Rys. 1. Budowa pojedynczego neuronu²

4. Matematyczny model sztucznego neuronu i sieci neuronowej

Na podstawie obserwacji mechanizmów zachodzących w naturalnej sieci neuronów, opracowano matematyczne koncepcje sztucznych sieci neuronowych (SSN). Podobnie jak w naturalnych sieciach, składają się one z elementarnych komórek - neuronów. Zdając sobie sprawę z braku możliwości dokładnego odwzorowania naturalnych układów i budowy, opracowano model sztucznego neuronu (patrz rys. 2. Jego działanie można przedstawić wg następujących kroków:

- Krok 1:** Do neuronu docierają sygnały (na wejścia neuronu).
- Krok 2:** Każdy sygnał ma swoją wagę (efektywność synapsy).
- Krok 3:** W neuronie obliczana jest ważona suma wejść i odejmowana wartość progowa,
- Krok 4:** Wynik sumy ważonej wprowadzany jest jako argument funkcji aktywacji, a wynik funkcji jest wprowadzany na wyjście neuronu.

² Tadeusiewicz R.: *Elementarne wprowadzenie do techniki sieci neuronowych z przykładowymi programami*, Akademicka Oficyna Wydawnicza, Warszawa 1998, str. 23.



Rys. 2. Schemat budowy sztucznego neuronu

Neurony traktować można jako elementarne procesory o następujących właściwościach:

- Każdy neuron otrzymuje wiele sygnałów wejściowych i wyznacza na ich podstawie swoją „odповідź”, to znaczy jeden sygnał wyjściowy.
- Z każdym oddzielnym wejściem neuronu związany jest parametr w nazwany wagą odpowiadającą stopniowi ważności informacji.
- Sygnał wchodzący określonym wejściem jest najpierw przemnażany przez wagę danego wejścia, w związku z czym w dalszych obliczeniach uczestniczy już w formie zmodyfikowanej: wzmacnionej ($w > 1$) lub stłumionej ($w < 1$).
- Sygnały wejściowe po przemnożeniu przez odpowiednie wagi są w neuronie sumowane, dając pewien pomocniczy sygnał wewnętrzny, nazywany czasem łącznym pobudzeniem neuronu (net value).
- Do tak utworzonej sumy sygnałów dodaje się pewien dodatkowy składnik, niezależny od sygnałów wejściowych, nazwany progiem (bias).
- Suma przemnożonych przez wagi sygnałów wejściowych z dodanym (ewentualnie) progiem jest traktowana bezpośrednio jako sygnał wyjściowy neuronu. W wielu typach sieci tak zwanych liniowych to wystarcza np. w sieciach ADALINE (ADaptive LINEar). Natomiast w sieciach o bogatszych możliwościach np. MLP (MultiLayer Perceptron) sygnał wyjściowy neuronu obliczany jest za pomocą pewnej nieliniowej zależności między łącznym pobudzeniem a sygnałem wyjściowym.
- Zależność sygnału wyjściowego od łącznego pobudzenia, zwana charakterystyką neuronu (transfer function) pozwala w każdej chwili jedno-

znacznie określić sygnał wyjściowy neuronu, jeśli znane są jego sygnały wejściowe³.

Jak widać na rys. 2, model sztucznego neuronu składa się z dwóch bloków: bloku sumowania Σ i bloku aktywacji $f(e)$. W pewnym przybliżeniu blok sumowania odpowiada biologicznemu ciału komórki, w której realizowane jest algebraiczne sumowanie ważonych sygnałów wejściowych oraz generowany jest sygnał wyjściowy e , który może być traktowany jako potencjał membranowy komórki.

Do zapisu zależności pomiędzy wejściem i wyjściem neuronu wygodnie jest stosować notację wektorową. Zestaw sygnałów wejściowych można przedstawić jako następujący wektor: $X = \langle x_1, x_2, \dots, x_n \rangle^T$, gdzie x_1, x_2, \dots, x_n – sygnały wejściowe, T jest symbolem transpozycji. Podobnie można przedstawić wektor współczynników wagowych, reprezentujących wiedzę, którą neuron nabywa w procesie uczenia: $W = \langle w_1, w_2, \dots, w_n \rangle^T$ gdzie w_1, w_2, \dots, w_n – odpowiednie wagi sygnałów wejściowych.

Powyższe wektory można interpretować jako punkty w n - wymiarowej przestrzeni. Wszystkie możliwe punkty, reprezentujące wejście i współczynniki wagowe tworzą odpowiednio przestrzeń wejść i przestrzeń wag. Matematyczny opis neuronu polega na zbudowaniu matematycznej funkcji wiążącej sygnały wejściowe i współczynniki wagowe w taki sposób, by możliwe było obliczenie sygnału wyjściowego w neuronie. Sygnał wyjściowy związany jest z sygnałem wejściowym poprzez funkcję f określającą działanie neuronu: $y = f(x_1, x_2, \dots, x_n)$.

W przypadku wybrania do eksperymentów neuronu liniowego zależność między wejściami ($x_i, i=1, 2, \dots, n$) a wyjściem y ma postać: $y = \sum_{i=1}^n w_i x_i$ lub korzystając z oznaczeń wektorowych: $y = W^T X$.

Element opisany powyższym równaniem liniowym ma ciekawą własność wyróżniania tych wektorów wejściowych, które są podobne do jego wektora wag. Sygnał wyjściowy y będzie tym większy, im bardziej położenie wektora wejściowego X w przestrzeni wejść będzie przypominać położenie wektora wag W w przestrzeni wag.

³ Tadeusiewicz R.: *Elementarne wprowadzenie do techniki sieci neuronowych z przykładowymi programami*, Akademicka Oficyna Wydawnicza, Warszawa 1998, str. 25.

W neuronie nieliniowym informacja wejściowa przetwarzana jest w sposób, który można opisać równaniem: $y = f(e)$, gdzie f jest wybraną funkcją nieliniową, a sygnał e odpowiada najczęściej łącznemu pobudzeniu neuronu, zgodnym z formułą przyjmowaną dla modelu liniowego: $e = \sum_{i=1}^n w_i x_i$ lub uzupełnioną dodatkową stałą składnik θ , tzw. prog $e = \sum_{i=1}^n w_i x_i \pm \theta$. W wielu praktycznych zastosowaniach próg jest odejmowany od powyższej sumy.

Jeżeli

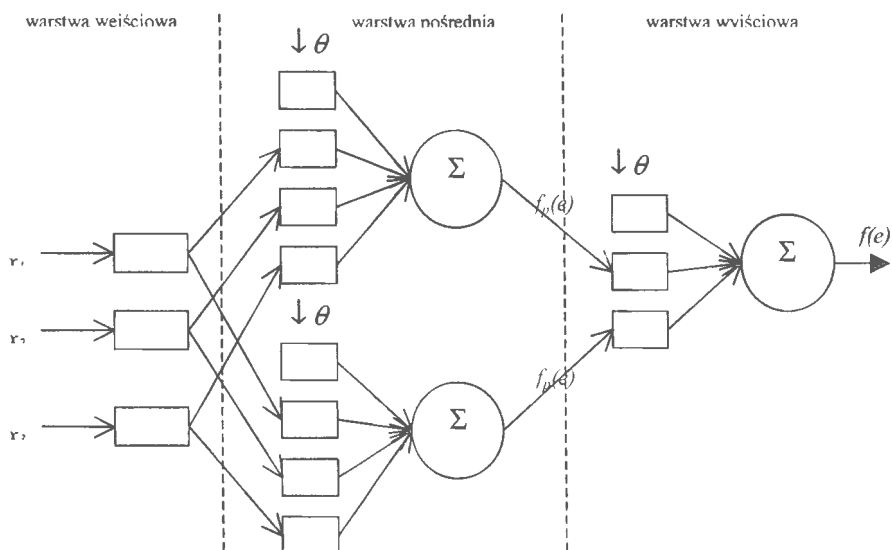
$$\sum_{i=1}^n w_i x_i \begin{cases} \geq \theta & \text{to dany wektor } X \text{ należy do klasy ponad progiem (klasa } A \text{)}, \\ < \theta & \text{to dany wektor } X \text{ należy do klasy pod progiem (klasa } B \text{)}. \end{cases}$$

Neuron można traktować jako element sieci przetwarzającej informacje. O specyficznych właściwościach tego elementu decyduje funkcja f określająca nieliniowy związek pomiędzy wypadkowym sygnałem wejściowym e , a odpowiedzią y . Najbardziej klasyczną funkcją nieliniową f używaną w sieciach neuronowych jest funkcja progowa oraz funkcja sigmoidalna.

Jednokierunkowa, wielowarstwowa sztuczna sieć neuronowa powstaje przez połączenie opisanych wyżej sztucznych neuronów. Stosowana jest przy tym zasada łączenia każdego neuronu warstwy poprzedniej z każdym neuronem warstwy następnej. Powstające w ten sposób warstwy można podzielić na: wejściową, pośredniczą i wyjściową, jak to przedstawiamy na rys. 3. Sieć nazywa się jedнокierunkową, gdyż informacja płynie w niej zawsze w jednym kierunku od wejścia do wyjścia, lewa na prawo.

Działanie sieci neuronowej można przedstawić w następujących krokach:

- Krok 1:** Dane wejściowe są wprowadzane na wejścia warstwy wejściowej.
- Krok 2:** Dane są propagowane na wejścia następnych warstw, aż do warstwy wyjściowej.
- Krok 3:** Wartości otrzymane w ostatniej, wyjściowej warstwie sieci neuronowej traktuje się jako wyjścia sieci.



Rys. 3. Schemat budowy wielowarstwowej, nieliniowej sztucznej sieci neuronowej

5. Uczenie sieci

Uczenie sieci z tzw. „nauczycielem”, polega na podawaniu do sieci przykładowych zadań wraz z rozwiązaniami. Proces ten polega na naśladowaniu sposobu działania nauczyciela, czyli zbioru przykładów (ciągu uczącego) na podstawie, których sieć się uczy.

Założmy, że rozważana sieć ma n wejść i k wyjść. Wiąże się to oczywiście z tym, że w sieci takiej jest n neuronów w warstwie wejściowej i k neuronów w warstwie wyjściowej. Przy takich założeniach ciąg uczący dla nauki pojedynczego neuronu ma postać:

$$U = \{ \langle X^{(1)}, z^{(1)} \rangle, \langle X^{(2)}, z^{(2)} \rangle, \dots, \langle X^{(N)}, z^{(N)} \rangle \}, \quad (1)$$

gdzie: N jest liczbą wektorów w ciągu uczącym;

$\langle X^{(j)}, z^{(j)} \rangle$ są parami danych pokazywanych na wejściu i wyjściu sieci (neuronu) w j -tym kroku procesu uczenia;

$X^{(j)}$ jest n -wymiarowym wektorem danych wejściowych w j -tym kroku;

$z^{(j)}$ jest wymaganą odpowiedzią neuronu w j -tym kroku (patrz rys.4).

W przypadku całej sieci wielowarstwowej ciąg uczący ma postać:

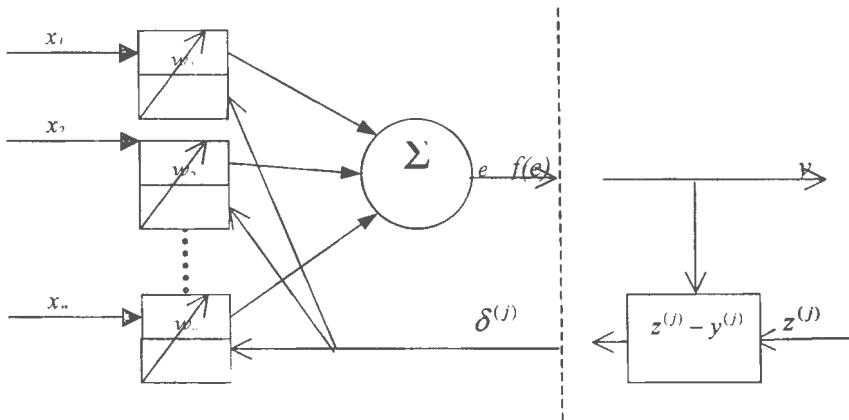
$$U = \{ \langle X^{(1)}, Z^{(1)} \rangle, \langle X^{(2)}, Z^{(2)} \rangle, \dots, \langle X^{(N)}, Z^{(N)} \rangle \}, \quad (2)$$

gdzie:

$X^{(j)}$ jest n -wymiarowym wektorem danych wejściowych w j -tym kroku;

$Z^{(j)}$ jest k -wymiarowym wektorem stanowiącym wymaganą odpowiedź neuronów wyjściowej warstwy sieci w j -tym kroku.

Przy uczeniu bez nauczyciela (samouczenie sieci) nie podaje się prawidłowych odpowiedzi. Sieć musi sama wydobyć interesujące kategorie lub cechy ze zbioru danych wejściowych. Samouczenie sieci jest analogią do działania ludzkiego mózgu. Człowiek również ma zdolność do spontanicznego klasyfikowania napotkanych obiektów i zjawisk, a później do rozpoznawania nowych obiektów jako należących do jednej z wcześniej wyuczonych klas. Metoda samouczenia jest oparta na regule Hebba, która zakłada wzmocnienie połączeń między źródłami silnych sygnałów a neuronami, które na nie silnie reagują.



Rys. 4. Schemat uczenia nadzorowanego dla elementu sieci

W wyniku uczenia sygnały podobne do siebie będą rozpoznawane (wzmocniane) przez pewne neurony, zaś przez inne będą ignorowane (osłabiane). Inaczej uczenie wg reguły Hebba nazywa się uczeniem korelacyjnym (correlation learning), gdyż zmierza ona do takiego dopasowania wag, aby uzyskać jak największą

korelację między sygnałami wejściowymi, a zapanowanym w formie wartości wag „wzorcem” sygnału, na który określony neuron, ma reagować.⁴

Omawianie algorytmów uczenia z nauczycielem zaczniemy od pokazanie sposobu uczenia pojedynczego (mieliniowego) neuronu. Podstawowym wzorem, wykorzystywanym podczas uczenia sieci neuronowych jest wzór

$$\Delta w_i^{(j)} = \eta \delta^{(j)} \frac{df(e)}{de^{(j)}} x_i^{(j)} \quad (3)$$

gdzie:

$\Delta w_i^{(j)}$ - zmiana i -tego współczynnika w i -tym kroku.

η - współczynnik uczenia się sieci;

$\delta^{(j)}$ - błąd j -tego obiektu ciągu uczącego

$\frac{df(e)}{de^{(j)}}$ - pochodna funkcji aktywacji.

$x_i^{(j)}$ - sygnał na i -tym wejściu j -tego obiektu.

Z (3) wynika, że zmiana $\Delta w_i^{(j)}$ współczynnika wagi $w_i^{(j)}$ na i -tym wejściu neuronu po pokazaniu j -tego obiektu ciągu uczącego jest proporcjonalna do błędu $\delta^{(j)}$ popełnionego przez neuron na tym etapie procesu uczenia.

$$\delta^{(j)} = z^{(j)} - y^{(j)} \text{ gdzie oczywiście } y^{(j)} = f\left(\sum_{i=0}^n w_i^{(j)} x_i^{(j)}\right)$$

Wynika z tego, że jeśli neuron podczas uczenia nie popełnia błędów, jego wagi nie są zmieniane. W ten sposób proces uczenia zatrzymuje się samorzutnie po osiągnięciu sukcesu, gdy żadna z wag nie uległa zmianie w poprzednim kroku. Poprawka współczynnika wagi na określonym wejściu $\Delta w_i^{(j)}$ proporcjonalna jest także do wielkości sygnału na tym wejściu $x_i^{(j)}$, co powoduje, że zmianom i aktualizacji podczas uczenia podlegają wagi tylko tych wejść, które są aktywne.

⁴ Tadeusiewicz R.: Elementarne wprowadzenie do techniki sieci neuronowych z przykładowymi programami, Akademicka Oficyna Wydawnicza, Warszawa 1998, str. 37.

Współczynnik η nazywany jest współczynnikiem uczenia (learning rate) i wpływa na szybkość procesu uczenia. Aby łatwo można było obliczyć pochodną $\frac{df(e)}{de^{(j)}}$ przyjmuje się zwykle jako funkcję logistyczną $f(e)$:

$$y = f(e) = \frac{1}{1 + \exp(-\beta e)}, \text{ dla której } \frac{df(e)}{de^{(j)}} = y^{(j)}(1 - y^{(j)}).$$

W związku z tym ostatecznie otrzymujemy formułę uczenia dla pojedynczego neuronu i dla dowolnej sieci jednowarstwowej w postaci:

$$\Delta w_i^{(j)} = \eta(z^{(j)} - y^{(j)})(1 - y^{(j)})x_i^{(j)}y^{(j)}. \quad (4)$$

Opisany powyżej algorytm nie daje się zastosować dla sieci wielowarstwowej, gdyż w przypadku sieci wielowarstwowych nie można bezpośrednio określić wymaganych odpowiedzi sygnałów wyjściowych $z^{(j)}$ dla warstw pośrednich sieci, a tym samym określić błąd $\delta^{(j)}$. Dlatego w odniesieniu do neuronów z tzw. warstw pośrednich także stosuje się wzór:

$$\Delta w_i^{(m)(j)} = \eta \delta_m^{(j)} \frac{d\varphi(e)}{de_m^{(j)}} y_i^{(j)}, \quad (5)$$

gdzie m jest numerem rozważanego neuronu.

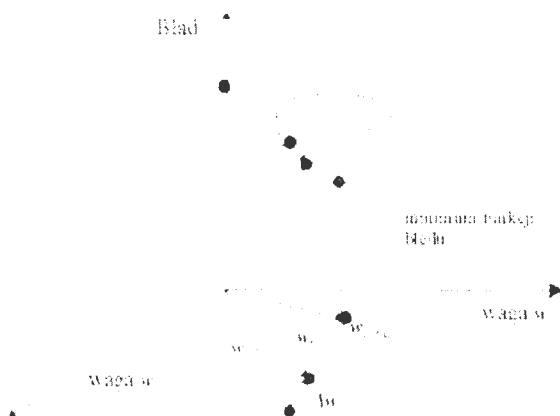
Algorytm wstecznej propagacji błędów jest podstawowym algorytmem uczenia nadzorowanego wielowarstwowych jednokierunkowych sieci neuronowych. Polega na takiej zmianie wag sygnałów wejściowych każdego neuronu w każdej warstwie, by wartość błędu dla kolejnych par uczących, zawartych w zbiorze uczącym była jak najmniejsza. Wykorzystuje on w tym celu metodę gradientową - najszybszego spadku (patrz rys.5).

Ogólny przebieg uczenia nadzorowanego sieci neuronowej przebiega według algorytmu przedstawionego na rys. 6.

Przebieg algorytmu dla wszystkich elementów ciągu uczącego nazywa się epoką. Rzutowane wstecznie błędy są mnożone przez te same współczynniki, przez które mnożone były sygnały przy propagowaniu sygnału w kierunku od wejścia do

wyjścia. Opisaną zasadę można zastosować do całej sieci, przy zachowaniu odpowiedniej kolejności obliczania błędów⁵.

Uczenie w sieci neuronowej polega więc na tym, że następuje w niej wielokrotnie odpowiednia zmiana współczynników wag reprezentujących wewnętrzną wiedzę neuronów w sieci. W wyniku tego błąd w sieci systematycznie maleje. Proces ten oryginalnie jest dość powolny – w praktyce czasem trzeba wykonać kilkaset milionów kroków uczenia, zanim sieć zacznie funkcjonować poprawnie, ale potem działa ona stosunkowo pewnie i niezawodnie.



Rys. 5. Istota metody zmiany wektora wag podczas uczenia sieci⁶

6. Pomiar KL ucznia z wykorzystaniem SSN

Sieci neuronowe charakteryzują się wieloma nietypowymi właściwościami, które w sposób istotny poprawiają możliwości prowadzenia badań i interpretacji ich wyników. Zasadniczą własnością sieci neuronowych jest ich zdolność do przyswajania wiedzy oraz umiejętność generalizacji, czyli uogólnienia tej wiedzy. Po-

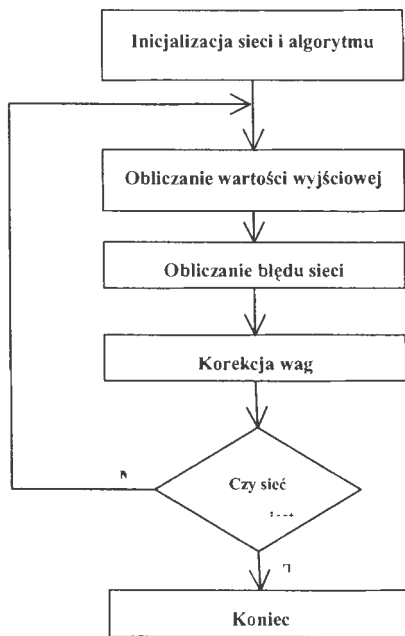
⁵ Tadeusiewicz R.: *Elementarne wprowadzenie do techniki sieci neuronowych z przykładowymi programami*, Akademicka Oficyna Wydawnicza, Warszawa 1998, str. 42.

⁶ Tadeusiewicz R.: *Elementarne wprowadzenie do techniki sieci neuronowych z przykładowymi programami*, Akademicka Oficyna Wydawnicza, Warszawa 1998, str. 17.

nadto sieci neuronowe wykazują tzw. tolerancję na nieciągłość, przypadkowe zaburzenia lub braki w zbiorze uczącym.

Nie ma konieczności precyzowania sposobu rozwiązania danego problemu, lecz wystarczy zgromadzić wystarczająco duży i reprezentatywny zbiór próbek (danych). Sieci neuronowe potrafią w naturalny sposób przetworzyć podaną w ten sposób informację i wyciągnąć z niej najistotniejsze elementy. Szczególnie ważne jest to w problemach klasyfikacji oraz wszelkiego rodzaju prognozach.

Celem artykułu jest zaproponowanie i weryfikacja własności sieci neuronowej ze wsteczną propagacją błędów w pomiarze kapitału ludzkiego ucznia szkoły ponagimnazjalnej na podstawie uzyskanych wyników z egzaminów zewnętrznych: maturalnego i gimnazjalnego.



Rys. 6. Schemat przebiegu algorytmu wstecznej propagacji błędów

Wyodrębniono następujące cele cząstkowe:

1. Zaproponowanie struktury sieci neuronowej oraz weryfikacja dokładności pomiaru w zależności od konfiguracji sieci.

2. Wyznaczenie poziomu kapitału ludzkiego $V(KL)$ wykorzystując wyznaczony typ i strukturę sieci.

Specyfika wykorzystania sieci neuronowych do tego celu zastosowań niemal narzuca sieci ze wsteczną propagacją błędów. Sieci tego typu z definicji umożliwiają bowiem uczenie z nauczycielem i mogą zawierać w tym przypadku jedno z dwóch możliwych wyjść – prawda lub fałsz. Inne typy sieci, uczące się bez nauczyciela (uczenie nienadzorowane), ze względu na swoją specyfikę (nieznane są klasy, do których należą mierzone wektory cechy), są rzadko wykorzystywane w tego typu zastosowaniach. Na wejście sieci podawane są zazwyczaj sekwencyjnie kolejne dane, dla których przeprowadzany jest proces rozpoznawania danych jako wektor wejściowy. Składa się on z kolumny danych liczbowych reprezentujących wyniki uczniów na egzaminie gimnazjalnym w poszczególnych latach.

7. Model SSN do pomiaru KL

W wyniku naszych badań proponujemy model sztucznej sieci neuronowej do obliczenia poziomu kapitału ludzkiego. W obliczeniach zastosowana jest sieć typu MLP (perceptron) – jest siecią wielowarstwową, nadzorowaną i jednokierunkową (patrz rys.7).

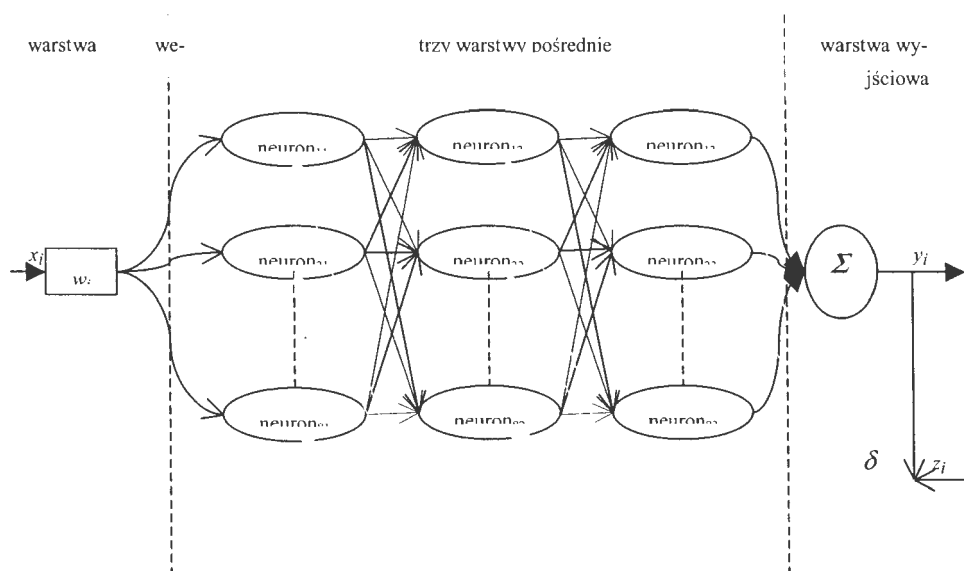
Sieć składa się z warstwy wejściowej, w której podawane są sygnały odpowiadające danym wejściowym x_i (wyniki gimnazjalne $v(KL, t_0)$), trzech warstw pośrednich, w których sieć trenowana jest na ośmiu neuronach w każdej warstwie oraz z warstwy wyjściowej y_i (poziom kapitału ludzkiego ucznia $V(KL)$). W modelu tym zastawaliśmy w pierwszej warstwie nieliniową funkcję aktywacji (sigmoidealną), w pozostałych warstwach funkcje liniowe. Sieć trenowana jest „z nauczycielem” tzn. otrzymany wynik z sieci y_i porównywany jest z wynikiem oczekiwanym z_i (wynikiem maturalnym $v(KL, t_k)$). Zadaniem sztucznej sieci neuronowej jest wyliczenie $V(KL)$ oraz narysowanie krzywej regresji - zależności wyników maturalnych od gimnazjalnych.

W procesie uczenia sieci wykorzystano metodę wstecznej propagacji błędów, której algorytm opiera się na sprowadzeniu wartości funkcji błędu sieci poniżej pewnego, założonego minimum.

Na podstawie przygotowanych wcześniej danych (wyniki uczniów z egzaminów gimnazjalnych i maturalnych odczytane na znormalizowanej skali staniowej) sieć uczy się prognozowania wyników oczekiwanych. Następnie obliczana

jest edukacyjna wartość dodana EWD, czyli różnica pomiędzy wynikiem uzyskanym przez ucznia a wynikiem oczekiwanym. Na tej podstawie będziemy mogli oszacować przyrost poziom kapitału ludzkiego ucznia $V(KL)$.

Podamy teraz formalny opis Metody SSN .



Rys. 7. Model sieci neuronowej do obliczania $V(KL)$

Metoda SSN– wykorzystanie sztucznych sieci neuronowych.

Krok 1: Dla przyjętych parametrów sieci uczymy ją na przykładach (uczniach) z zadanego ciągu uczącego.

Krok 2: Za pomocą tak nauczonej sieci dla każdego ucznia z badanej populacji wyznaczamy wartość poziomu jego kapitału ludzkiego $\hat{v} = (KL, t_k)$

Krok 3: Dla każdego ucznia z badanej populacji wyznaczamy wkład szkoły równy

$$V(KL) = v(KL, t_k) - \hat{v}(KL, t_k). \quad (6)$$

Krok 4: Za pomocą Tabeli 1 interpretujemy wyniki uzyskane w Kroku 3.

Wnioski

Nasze badania wykazały dużą przydatność sieci neuronowych do opracowywania i analizy szacowania poziomu kapitału ludzkiego ucznia. Okazało się, że elastyczność sieci neuronowych umożliwia poprawną estymację nieznanymi zależnościami funkcyjnymi. Techniki statystyczne w rozważanych problemach dawały nieco gorsze rezultaty. Istnieje zatem w praktyce możliwość wykorzystywania sztucznych sieci neuronowych do analizy zbiorów danych. Taką analizę można zawsze weryfikować lub doprecyzować przy pomocy standardowych metod statystycznych.

Oczywiście, nawet najbardziej dokładna predykcja wyników egzaminu wyższego szczebla nie rozwiązuje w pełni tak złożonego problemu jak szacowanie poziomu kapitału ludzkiego ucznia. Daje ona jednak podstawy do dalszych badań i analiz tego problemu. Ponadto, zdaniem autorki, dalsze badania nad Wirtualną Taśmą Produkcyjną oraz sztucznymi sieciami neuronowymi, powinny dać nowe narzędzia umożliwiające coraz to dokładniejszą i bardziej adekwatne szacowania poziomu kapitału ludzkiego, pojedynczej osoby (ucznia, pracownika itp.), jak też zespołu, grupy czy firmy.

Literatura

- [1]. Bishop C.M. (1995): *Neural Networks for Pattern Recognition*. Oxford, Clarendon Press.
- [2]. Czajka M. (2005): *Matlab ćwiczenia*. Helion.
- [3]. Duch W., Korbicz J., Rutkowski L., Tadeusiewicz R. (2000): *Sieci neuronowe. Biocybernetyka i inżynieria biomedyczna*, tom 6. Exit. Warszawa.
- [4]. Geman S., Bienenstock E., Doursat R. (1992): Neural networks and the bias – variance dilemma. *Neural Computation*.
- [5]. Golden R.M. (1996): *Mathematical Methods for Neural Network Analysis and Design*. Cambridge, MA, MIT Press.
- [6]. Hand D., Mannila H., Smith P. (2005): *Eksploracja danych*. WN-T, Warszawa.
- [7]. Korbicz J., Obuchowicz A., Uciński D. (1994): *Sztuczne sieci neuronowe. Podstawy i zastosowania*. Akademicka Oficyna Wydawnicza PLJ, Warszawa.
- [8]. Larose D. T. (2006): *Odkrywanie wiedzy z danych. Wprowadzenie do eksploracji danych*. PWN, Warszawa.

ISBN 9788389475220