Piotr Boniecki
Instytut Inżynierii Rolniczej
Akademia Rolnicza w Poznaniu

LINIOWE SIECI NEURONOWE A METODY ANALIZY REGRESJI W ASPEKCIE ICH WYKORZYSTANIA W INŻYNIERII ROLNICZEJ

Streszczenie

Nieustanne dążenie badaczy do pełniejszego rozumienia i wyjaśnienia praw rządzących przyrodą spowodowało, że rosnącego znaczenia nabierają poszukiwania nowych metod badawczych, coraz efektywniej wspomagających procesy poznawcze. Należą do nich niewątpliwie uzupełniające modele symulacyjne, tworzone dedukcyjnie na zbiorach przesłanek, wynikających z aktualnego stanu wiedzy naukowej. Techniki eksperymentu wirtualnego, wspomagające proces badania złożonych systemów empirycznych, powinny znajdować zastosowanie praktyczne również w dyscyplinie naukowej, jaką jest inżynieria rolnicza. Dynamiczny rozwój technik informatycznych spowodował pojawienie się zupełnie nowych możliwości obliczeniowych, bazują-cych na wzorcach pochodzących bezpośrednio z obserwacji procesów naturalnych, a w szczególności pracy mózgu. Kluczową rolę spełniają tu metody sztucznych sieci neuronowych, stanowiące w wielu przypadkach modele ekwiwalentne (a często znacznie rozszerzające potencjalne widmo zastosowań) w stosunku do tradycyjnych metod statystycznych.

Słowa kluczowe: sieci neuronowe, metody analizy regresji

Wprowadzenie

Podczas realizacji empirycznych badań prowadzonych nad złożonymi systemami inżynierii rolniczej często zdarza się, że nie dysponujemy jawnym, sformalizowanym modelem matematycznym, precyzyjnie opisującym dane zagadnienie. Oznacza to, że posiadamy jedynie wyniki badań, podczas gdy nasza wiedza o strukturze badanego problemu jest zazwyczaj fragmentaryczna, a zatem w dużym stopniu niepełna. Wiąże się to niewątpliwie ze wspomnianą już złożoną, często losową strukturą, jaką reprezentuje większość systemów inżynierii rolniczej. Dotychczas w takich wypadkach stosowano metody analizy regresji, starając się określić stopień stochastycznej zależności zachodzącej pomiędzy poszczególnymi zmiennymi losowymi, co następnie stanowiło podstawę do tworzenia różnych modeli empirycznych.

Rozwój technologii informatycznych spowodował pojawienie się nowych możliwości analitycznych, bazujących na obserwacjach procesów zachodzących w naturze, a w szczególności na wnioskach płynących z badań naukowych dotyczących budowy oraz pracy mózgu. Pojawiające się metody oraz techniki, posiadające znamiona sztucznej inteligencji, pozwalają na tworzenie modeli symulacyjnych, które realizują postawione zadania w oparciu o wzorce zaczerpnięte bezpośrednio z obserwacji przyrody. W technikach optymalizacyjnych wykorzystuje się coraz częściej procedury oparte na algorytmach ewolucyjnych, natomiast przetwarzanie informacji realizowane jest często w sposób równoległy oraz rozproszony. Należy podkreślić, że sztuczne sieci neuronowe potrafią operować zarówno na zbiorach danych numerycznych, pochodzących np. z badań empirycznych, jak również na zaszumionych zbiorach rozmytych, tak charakterystycznych dla postrzegania ludzkiego umysłu. Nominalne zmienne rozmyte opisujące w sposób naturalny zjawiska i procesy zachodzące w przyrodzie, znajdują coraz częściej zastosowanie w prowadzonych obecnie badaniach naukowych. Niestety, do tej pory w dyscyplinie inżynieria rolnicza, nie wykorzystywano w pełni możliwości jakie daje stosowanie omawianych metod sztucznej inteligencji. W szczególności, obiecujące rezultaty przynosi wykorzystanie takich aspektów sztucznych sieci neuronowych, jak zdolność do rozwiązywania zagadnień regresyjnych oraz klasyfikacyjnych.

Stosunkowo najprostsze w budowie i eksploatacji są liniowe sieci neuronowe. Pozwalają one na wykonanie zbliżonych zadań do tych, jakie stawiane są zazwyczaj w trakcie „klasycznej obróbki” danych empirycznych. Dzieje się tak dlatego, ponieważ modelowanie liniowe w istocie polega na aproksymacji funkcji dyskryminującej (sieci neuronowe) lub na opisie charakteru stochastycznej zależności między zmiennymi (analiza regresji), realizowanego za pomocą modelu w formie hiperpłaszczyzny [Papoulis 1972]. Celem pracy jest wskazanie podobieństw oraz różnic jakie reprezentują liniowe modele neuronowe oraz tradycyjne metody analizy regresji. Dodatkowo podjęto próbę oszacowanie zakresu aplikacji tych modeli w badaniach empirycznych prowadzonych w dyscyplinie inżynieria rolnicza.

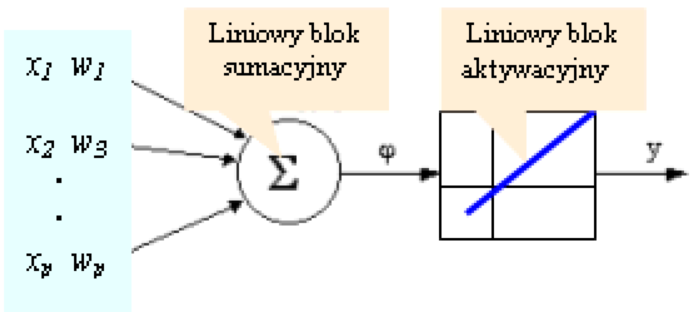
Liniowe sieci neuronowe

W badaniach naukowych prowadzonych w dyscyplinie inżynieria rolnicza, często można napotkać problem wyglądający na trudny i nieliniowy, jednak dający się stosunkowo łatwo rozwiązać w oparciu o szeroko rozumiane metody oraz techniki liniowe [Tadeusiewicz 1993].

Faktem jest, że popularność sztucznych sieci neuronowych wynika głównie z możliwości modelowania zagadnień nieliniowych czyli problemów opisywanych modelem regresji krzywoliniowej. Nie oznacza to jednak by należało zapomnieć w procesie analizy danych empirycznych o stosunkowo prostych modelach liniowych. Ogólna zasada stosowana w nauce głosi, że w przypadku gdy istnieje możliwość wyboru pomiędzy modelem prostym i bardziej złożonym, należy zawsze preferować model prostszy (o ile oczywiście model złożony nie dopasowuje się znacząco lepiej do posiadanych danych). Warto zauważyć, że najprostszym modelem aproksymującym zaobserwowaną zależność empiryczną opisaną matematycznie, jest właśnie model liniowy. W modelu takim funkcją dopasowywaną do posiadanych danych jest zawsze hiperpłaszczyzna, a zadanie w istocie polega na znalezieniu właściwej orientacji tej hiperpłaszczyzny w przestrzeni zdarzeń. W liniowych zagadnieniach klasyfikacyjnych zorientowana jest ona w taki sposób, aby dobrze odzwierciedlała generalną tendencję reprezentowaną przez dane wejściowe, które podlegają analizie. Warto pamiętać, że model liniowy jest bardzo często dobrym odniesieniem, z którym można porównywać bardziej złożone struktury.

Mając powyższe na uwadze omówiony zostanie jedynie wybrany aspekt aplikacji sieci neuronowych, jako ekwiwalentnego narzędzia do badania zagadnień, dających się w szczególności modelować w oparciu o liniową analizę regresji.

Matematyczny opis sztucznego neuronu liniowego opracowali w 1943 roku Warren McCulloch i Walter Pitts [Tadeusiewicz 1993; Korbicz i in. 1994]. Założyli oni, że sztuczny neuron jest uproszczonym modelem matematycznym symulującym działanie biologicznej komórki nerwowej, realizującej złożony proces przetwarzania informacji.



Rys. 1. *Struktura sztucznego neuronu liniowego*
Fig. 1. *Structure of artificial linear neuron*

Jego zasadniczym celem jest przetworzenie sygnału wejściowego (dostarczanego w postaci wartości $x_1, x_2, ..., x_i$) w wartość wyjściową y, stanowiącą rezultat „działania" neuronu. Zarówno wartości wprowadzane na wejścia neuronu, jak i uzyskiwana na jego wyjściu wartość wyjściowa, mają postać liczb rzeczywistych. Z każdym i-tym wejściem neuronu związany jest współczynnik w_i nazwany wagą (synaptyczną). Prezentowane współczynniki wagowe w_i są podstawowymi, adaptacyjnymi parametrami wpływającymi na sposób funkcjonowania sztucznego neuronu. Przetwarzanie informacji wykonywane jest w dwóch etapach. Pierwszy z nich sprowadza się do realizacji procesu agregacji ważonej danych wejściowych x_i. Ma on na celu przetworzenie wektora wejściowego $\mathbf{x} = (x_1, x_2, ..., x_i)$ w pojedynczą wartość liczbową (tzw. zagregowaną wartość wejściową φ, określaną też w literaturze mianem potencjału postsynaptycznego lub membranowego). Druga faza to z kolei realizacja procesu wyznaczenia wartości wyjściowej y neuronu. Elementem odpowiedzialnym za wykonanie tej czynności jest funkcja aktywacji, stanowiąca integralną składową tzw. bloku aktywacji. Jest to w istocie funkcja przekształcająca wyznaczoną wcześniej zagregowaną wartość wejściową φ w wartość wyjściową neuronu.

Zgodnie z tym modelem wzór na potencjał membranowy φ (*PSP- **P**ost-**S**ynaptic **P**otential*) ma następującą postać:

$$\varphi = \sum_{i=1}^{P} x_i w_i \tag{1}$$

gdzie:

x_i – wartość i-tego sygnału wejściowego,
w_i – wartość i-tej wagi,
P – liczba wejść,
φ – potencjał membranowy (*PSP*),
$\sum_{i=1}^{P} x_i w_i = 0$ – równanie hiperpłaszczyzny (tzw. liniowa granica decyzyjna).

Otrzymana ze wzoru (1) wartość φ poddawana jest następnie przetwarzaniu przez tzw. blok aktywacyjny**,** stanowiąc argument dla przyjętej funkcji dyskryminacyjnej. W omawianym przypadku (model liniowy) jest to funkcja liniowa, więc sygnał wyjściowy przybiera postać następującą:

$$y = k\varphi = k\sum_{i=1}^{P} x_i w_i \tag{2}$$

gdzie:

y – wartość wyjściową neuronu,
k – stały współczynnik.

Ogólnie można powiedzieć, że rozważany neuron liniowy jest w istocie operatorem iloczynu skalarnego wektora wag oraz wektora wejściowego. Zapisując więc (1) w formie składowych w.w. wektorów, mamy:

$$y = k\sum_{i=1}^{P} x_i w_i = k \cdot |\vec{x}| \cdot |\vec{w}| \cdot \cos(\vec{x}, \vec{w}) \tag{3}$$

Warto zauważyć, że „model liniowy” jest pojęciem jednoznacznym. Dla danego zbioru danych empirycznych istnieje tylko jeden optymalny model liniowy i zadanie jego wyznaczenia (poprzez minimalizację błędu) jest zadaniem dobrze określonym, stąd można je stosunkowo łatwo rozwiązać.

Proces uczenia liniowej sieci neuronowej

Proces uczenia sieci neuronowej (składającej się z neuronów liniowych) sprowadza się do minimalizacji wielkości kąta zawartego pomiędzy wektorem wejść i wektorem wag (optymalizacji podlega iloczyn skalarny wektora wejść i wektora wag). Uzyskuje się to przez rekurencyjne „dostrojenie” wektora wag $\mathbf{w}[w_1, w_2 .....w_P]$, którego składowe modyfikowane są w kolejnych cyklach (tzw. epokach) procesu uczenia, do wektora wejść stanowiącego wzorzec. W trakcie ogólnego procesu uczenia nadzorowanego (uczenie z nauczycielem) wykorzystuje się zasadę działania liniowego sumatora ważonego, którego sygnał wyjściowy opisany jest wzorem [Duch, Tadeusiewicz, Korbicz i in. 2000]:

$$y(n) = \sum_{i=1}^{P} x_i(n) w_i = \mathbf{w}^T \mathbf{x}(n) \tag{4}$$

gdzie:

n – liczba cykli (sygnałów) uczących,
$\mathbf{w}^{\mathrm{T}}$ – transponowany wektor wag,
$\mathbf{x}$ – wektor wejść,
P – liczba wejść.

Traktując sygnał (4) jako estymatę pewnego sygnału wzorcowego (uczącego) oznaczonego przez $d(n)$ można określić błąd estymacji w następującej postaci:

$$\varepsilon(n) = d(n) - y(n) \tag{5}$$

gdzie:

$d(n)$ – sygnał wzorcowy (wektor uczący),
$y(n)$ – sygnał estymaty liniowego sumatora ważonego (4).

W trakcie procesu uczenia, wagi w_1, w_2,....w_P dobiera się w kolejnych cyklach tak, aby zminimalizować miarę błędu [Tadeusiewicz 1993], [Duch, Tadeusiewicz, Korbicz i in. 2000] (w sensie średniokwadratowym):

$$\Delta(\mathbf{w}) = E\left[\varepsilon(n)^2\right] = E\left[\left(d(n) - \mathbf{w}^T\mathbf{x}(n)\right)^2\right] \tag{6}$$

gdzie operator całkowy E (***E**xpectation* - wartość oczekiwana) jest postaci:

$$E[\cdot] = \int_{-\infty}^{+\infty} xf(x)dx \tag{7}$$

W szczególnym przypadku, adekwatnie dla klasy omawianego problemu, operator E upraszcza się do trywialnej operacji średniej arytmetycznej. Stąd równanie (6) można zredukować do postaci:

$$\Delta(\mathrm{w}) = Q(\mathrm{w}) = \frac{1}{2}\sum_{i=1}^{P}\left(d(n) - \mathrm{w}^T\mathrm{x}(n)\right)^2 \tag{8}$$

Funkcja $Q(\mathbf{w})$ jest powszechnie znana w teorii aproksymacji jako powierzchnia błędu średniokwadratowego. W omawianym przypadku jest ona jedynie adoptowana z teorii optymalizacji dla potrzeb realizacji procesu uczenia liniowej sieci neuronowej.

Można łatwo udowodnić [Tadeusiewicz 1993; Korbicz i in. 1994], że błąd średniokwadratowy estymacji (6) oraz (8) jest funkcja 2-go rzędu wektora wag **w**. Geometrycznie $Q(\mathbf{w})$ przedstawia hiperparaboloidę. W szczególnym przypadku (dla dwóch wejść) jest to paraboloida. Znalezienie optymalnych wartości składowych wektora wag **w** wyznaczającego ekstremum globalne funkcji błędu (w omawianym przypadku jest to minimum globalne), sprowadza się do wyznaczenia wektora gradientu ∇ funkcji $Q(\mathbf{w})$ i przyrównania otrzymanego rezultatu do 0:

$$\nabla Q = \frac{\partial Q}{\partial \mathbf{w}(n)} = \left[\frac{\partial Q}{\partial w_1}, \frac{\partial Q}{\partial w_2}, \ldots\ldots \frac{\partial Q}{\partial w_P}\right]^T = 0 \tag{9}$$

W oparciu o znaną z teorii programowania liniowego, gradientową metodę najszybszego spadku, uzyskuje się w kolejnych iteracjach (cyklach uczących) przybliżenia optymalnego wektora **w**. Uzyskuje się to dokonując w następujących po sobie krokach stosownej korekty w oparciu o wyliczoną wcześniej poprawkę $\Delta\mathbf{w}_n$. Występujące kolejne korekty wag $\mathbf{w}(n)$ powinny następować w kierunku przeciwnym do znaku składowych wektora gradientu:

$$\nabla Q(\mathbf{w}(n)) = \frac{\partial Q(\mathbf{w}(n))}{\partial \mathbf{w}(n)} = \left[\frac{\partial Q(\mathbf{w}(n))}{\partial w_1(n)}, \frac{\partial Q(\mathbf{w}(n))}{\partial w_2(n)}, \ldots\ldots \frac{\partial Q(\mathbf{w}(n))}{\partial w_P(n)}\right]^T \tag{10}$$

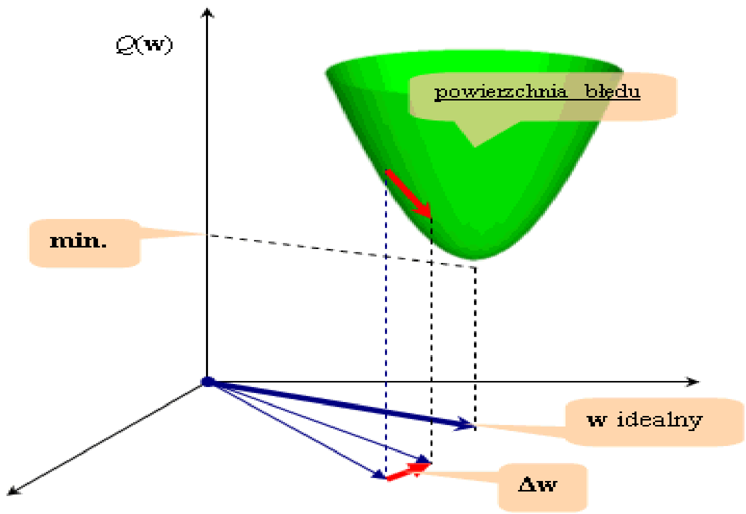
Zgodnie z powyższym dokonując stosownych przekształceń, algorytm najszybszego spadku można zapisać następująco [Tadeusiewicz 1993]:

$$\Delta \mathbf{w}_n = \mathbf{w}(n) - \mathbf{w}(n+1) = \frac{1}{2}\eta \frac{\partial Q(\mathbf{w}(n))}{\partial \mathbf{w}(n)} \tag{11}$$

gdzie:

$\eta > 0$ – wielkość kroku iteracyjnego (kroku korekcji),
n – ilość cykli (sygnałów) uczących.

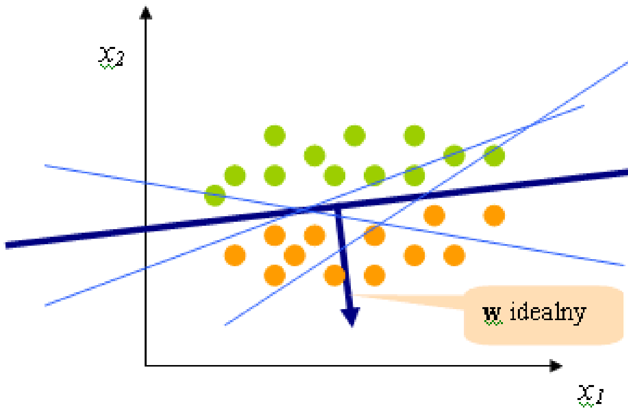
Proces uczenia sieci posiadającej dwa wejścia przedstawia rysunek 2.



Rys. 2. Interpretacja graficzna procesu uczenia liniowej sieci neuronowej
Fig. 2. Graphic interpretation of learnedly of neural linear network process

Proces uczenia jednokierunkowej sieci neuronowej (w tym również sieci liniowej) sprowadza się do rekurencyjnego „dostrojenia” wektora wag (parametry sieci) do wektora wejściowego, reprezentującego dane uczące stanowiące wzorzec. W istocie sieć działa jak liniowy klasyfikator dokonujący separacji liniowej na zbiorze wejściowym. Oczywiście jest to możliwe tylko w przypadku, gdy zbiór uczący reprezentuje swoją strukturą dane empiryczne mające charakter liniowy. Po pozytywnym zakończeniu procesu uczenia orientacja wektora wag sieci odpowiada sytuacji, w której hiperpłaszyzna reprezentowana przez „dostrojony" wektor wag separuje liniowo przestrzeń sygnałów podanych na wejściu sieci neuronowej.

Dla dwóch wejść procedurę tę można przedstawić na poglądowym rysunku:



Rys. 3. Neuron liniowy jako klasyfikator liniowy
Fig. 3. Linear neuron as linear classifier

W zagadnieniach klasyfikacyjnych prosta, której wektor normalny **w** uznano za idealny (dla wielu wymiarów będzie to hiperpłaszczyzna) jest umieszczana w takiej pozycji, aby oddzielać od siebie dwie rozważane klasy. Działa ona wtedy jako liniowa funkcja dyskryminująca (separująca).

Niewątpliwą zaletą modelowania liniowego w oparciu o sieć liniową jest możliwość przeprowadzenia efektywnej optymalizacji globalnej takiego modelu (co pozwala na szybkie i pewne wyznaczenie jego parametrów). Podstawową wadą jest jednak częsty brak zgodności z rzeczywistym charakterem rozważanych problemów, a postulowaną w modelu formą zależności liniowej. W konsekwencji badając liniowe sieci neuronowe modelujące problemy regresyjne można zauważyć, że dopasowanie hiperpłaszczyzny do wszystkich subtelności zawartych w zgromadzonych danych wejściowych bywa często utrudnione. Dzieje się tak głównie ze względu na występowanie złożonych, często nieliniowych relacji zachodzących w strukturze danych empirycznych.

O liniowej analizie regresji

Jak wiadomo, zaproponowana przez Legedre'a metoda najmniejszych kwadratów wykorzystana jest również jako narzędzie do szacowania wielkości błędu w metodach regresyjnych, powszechnie stosowanych w procesie analizy danych empirycznych [Papoulis 1972]. W zagadnieniach regresji liniowej estymacyjna procedura najmniejszych kwadratów, w połączeniu z gradientowymi metodami optymalizacyjnymi, pozwala na określenie parametrów prostych regresji. Umożliwia to estymację stochastycznych zależności zachodzących pomiędzy zmiennymi losowymi, reprezentującymi wyniki uzyskanych pomiarów (oczywiście, jeżeli taka zależność istnieje).

Analiza regresji (oraz korelacji) umożliwia badanie wzajemnego wpływu stochastycznych czynników mierzalnych, a więc danych pozyskanych w procesie empirycznym. Wyznaczenie prostej regresji II rodzaju daje możliwość oszacowania tego wpływu. Regresja prostoliniowa ma istotne znaczenie w badaniach prowadzonych w dyscyplinie inżynieria rolnicza, ponieważ w wielu przypadkach wielowymiarowa zmienna losowa $(X_1,X_2.....)$ ma w przybliżeniu rozkład normalny. Z teorii rachunku prawdopodobieństwa wiadomo, że jeśli linie regresji I rodzaju są prostymi, co jest równoważne normalności rozkładu zmiennej losowej $(X_1,X_2.....)$, to pokrywają się one z liniami regresji II rodzaju, wyznaczonymi metodą najmniejszych kwadratów. Podobnie jak w poprzednim rozdziale, wygodnie jest posłużyć się modelem 2-wymiarowym, a następnie uogólnić rozważania na przestrzeń wielowymiarową.

Przyjmując, że linia regresji X_1 względem X_2 jest postaci [Platt 1974]:

$$y = ax + b \tag{12}$$

można stworzyć wyrażenie (z zastosowaniem metody najmniejszych kwadratów):

$$\Delta(a,b) = E(X_2 - aX_1 - b)^2 \tag{13}$$

Dążymy następnie do tego, aby tak dobrać „*a*" oraz „*b*" by wyrażenie (13) osiągało minimum, tzn:

$$\begin{cases} \dfrac{\partial \Delta(a,b)}{\partial a} = 0 \\ \dfrac{\partial \Delta(a,b)}{\partial b} = 0 \end{cases} \tag{14}$$

Różniczkując a następnie rozwiązując układ (14) względem „*a*” oraz „*b*”, można otrzymać:

$$a = \rho \frac{\sigma_{x_2}}{\sigma_{x_1}}; \qquad b = m_{x_2} - \rho \frac{\sigma_{x_2}}{\sigma_{x_1}} m_{x_1} \tag{15}$$

gdzie:

ρ – współczynnik korelacji,

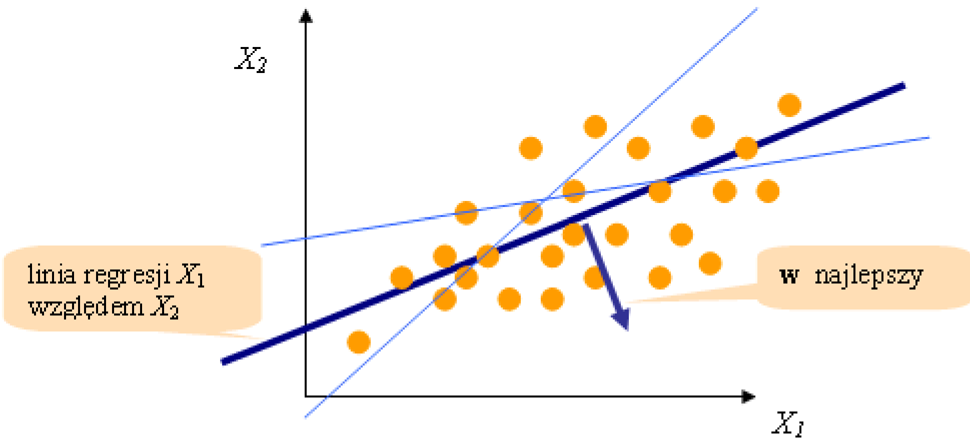
$\sigma_{x_1}, \sigma_{x_2}$ – odchylenia standardowe zmiennych losowych X_1 oraz X_2,

m_{x_1}, m_{x_2} – momenty 1-szego rzędu (wartości oczekiwane zmiennych losowych X_1 oraz X_2).

Wstawiając (15) do (12) otrzymujemy poszukiwaną linię regresji X_1 względem X_2 postaci:

$$y - m_{x_2} = \rho \frac{\sigma_{x_2}}{\sigma_{x_1}} \left(x - m_{x_1} \right) \tag{16}$$

Analogicznie można otrzymać linię regresji X_2 względem X_1.



Rys. 4. Linia regresji X_1 względem X_2
Fig. 4. The line of regress X_1 vrs. X_2

W zagadnieniach regresyjnych prosta uznana za najlepszą umieszczana jest w taki sposób, aby dobrze charakteryzowała generalną tendencję reprezentowaną przez dane empiryczne.

Zaletą modelowania liniowego jest możliwość przeprowadzenia optymalizacji globalnej takiego modelu (co pozwala na szybkie i pewne wyznaczenie jego parametrów). Niestety, jego podstawową wadą jest częsty brak zgodności z rzeczywistym (nieliniowym) charakterem rozważanych problemów, a postulowaną w modelu formą zależności liniowej.

Modele neuronowe oraz metody regresyjne w inżynierii rolniczej.

Modele regresyjne zajmują ugruntowaną pozycje w badaniach empirycznych prowadzonych w rolnictwie. Powstające w oparciu o metody statystyczne formuły empiryczne pozwalają na efektywny opis złożonych struktur, jakie reprezentują systemy inżynierii rolniczej. Konieczność probabilistycznej prezentacji omawianej klasy zagadnień wynika zarówno z braku posiadania pełnej wiedzy naukowej, jak również z faktu występowania losowych zakłóceń, towarzyszących realizacji samego procesu poznawczego. Przykładowe aplikacje zastosowania modeli statystycznych w dyscyplinie inżynieria rolnicza to m.in.:

– analiza struktury zbiorów płodów rolnych w celu dokonania właściwych prognoz;
– badanie zależności statystycznych pomiędzy parametrami pracy silników spalinowych;
– testowanie hipotez statystycznych, np.: badanie wpływu zużycia paliwa na poziom zbiorów płodów rolnych;
– szacowanie parametrów z populacji na podstawie losowo wybranej próby.

Można też wskazać wiele obszarów omawianej dyscypliny, w których zastosowanie sieci neuronowych wydaje się być szczególnie uzasadnione. Wymienię jedynie kilka przykładowych zastosowań:

– w procesie predykcji – np. prognozowanie poziomu zbiorów płodów rolnych na podstawie wyników z lat ubiegłych;
– do aproksymacji przestrzennej (znamy wartości parametrów (fizycznych, biologicznych itd.) w kilku punktach w terenie i pytamy, jakie wartości będą pomiędzy nimi) – np. analiza geologiczna gruntów itp.;
– do rozpoznawania kategorii (klasyczny problem klasyfikacyjny) – np. na podstawie charakterystyki cech owada, sieć neuronowa określa, do jakiej kategorii entomologicznej ten obiekt należy;
– do rozpoznawania obrazu – np. do sterowania mobilnym robotem wspomagającym pracę na roli;
– do filtrowanie sygnałów (obróbka danych zaszumionych) – np. analiza danych meteorologicznych.

Należy jednak zauważyć, że zastosowania sieci neuronowe posiadają również pewne ograniczenia. Problemem sieci jest niemożność uzyskania z ich pomocą ścisłej zależności matematycznej, występującej między wejściem a wyjściem sieci neuronowych. Dzieje się tak, ponieważ sztuczne sieci neuronowe są ze swej natury narzędziem o charakterze probabilistycznym. Faktem jest, że sieć tworzy neuronowy model badanego zjawiska, ale nie można precyzyjnie określić struktury tego modelu – znana jest jedynie topologia nauczonej sieci. Trudno ponadto dowieść ściśle, że nauczona sieć rzeczywiście potrafi sobie radzić z zadanym problemem. Istota procesu uczenia sieci neuronowej determinuje często stochastyczny charakter wyników podawanych na wyjściu przez tą sieć.

Uwagi końcowe

Nasuwa się oczywiste pytanie, które metody są lepsze? Należy przede wszystkim zauważyć, że obojętnie jaka metoda analizy danych, czy to statystyczna czy oparta na sieciach neuronowych, nie wytworzy informacji, jeśli nie będzie ona zawarta w uzyskanych danych pomiarowych. Stąd o wartości opisanych wyżej modeli w sposób zasadniczy decyduje jakość danych empirycznych, pozyskanych w trakcie prowadzonych badań naukowych.

Omówione analogie metodyczne, sugerują możliwość wykorzystania numerycznych symulatorów liniowych sieci neuronowych, jako narzędzia alternatywnego, a często wręcz komplementarnego, dla powszechnie stosowanych klasycznych metod regresji liniowej. Dodatkowym argumentem na korzyść neuronowej analizy danych, wydaje się być prosta struktura kodu sieci neuronowej, co ułatwia programiście jej zaprojektowanie oraz ewentualną implementację w systemie informatycznym (np. w tworzonym systemie ekspertowym). Nie bez znaczenia pozostaje też łatwość praktycznego użycia dostępnych na rynku programowych symulatorów sieci neuronowych. Kolejnym argumentem na korzyść stosowania sieci neuronowych wydaje się być spostrzeżenie, że wiele zagadnień oraz problemów inżynierii rolniczej nie posiada sformalizowanych modeli empirycznych a tym bardziej zweryfikowanych, teoretycznych modeli matematycznych. W takiej sytuacji możliwość stosunkowo łatwego pozyskania wiedzy naukowej ukrytej w danych empirycznych, wydaje się być szczególnie atrakcyjną.

Panu Profesorowi Stanisławowi Pabisowi serdecznie dziękuję za życzliwość oraz wszechstronną pomoc okazaną w trakcie pisania tej pracy

Bibliografia

Tadeusiewicz, R. 1993. Sieci neuronowe. Warszawa, Akademicka Oficyna Wydawnicza.

Korbicz J. i inni. 1994. Sztuczne sieci neuronowe – podstawy i zastosowania . Warszawa, Akademicka Oficyna Wydawnicza PLJ.

Duch W., Tadeusiewicz R., Korbicz J. i inni. 2000. Sieci neuronowe. Warszawa, Akademicka Oficyna Wydawnicza EXIT.

Platt Cz. 1974. Problemy rachunku prawdopodobieństwa i statystyki matematycznej, Warszawa, PWN.

Papoulis A. 1972. Prawdopodobieństwo, zmienne losowe i procesy stochastyczne. Warszawa WNT.

LINEAR NEURAL NETWORKS VS. REGRESSION ANALYSIS METHODS IN THE ASPECT OF THEIR APPLICATIONS IN AGRICULTURAL ENGINEERING

Endless efforts made by researches in order to better understand and explain principles governing the nature, has caused that it is becoming of greater importance to seek new investigation methods, which play an increasingly more significant role in enhancing the cognitive processes. Such are, beyond all doubt, the supplementary simulation models, created by inference based on the sets of indications, resulting from the current status of knowledge. Virtual experimentation techniques, aiding the process of examining complex empirical systems, should be utilized practically, also in such domain as the agricultural engineering. Dynamic growth of IT techniques has brought completely new computing capacities, based on the examples originating directly from observation of natural processes, especially the function of brain. The methods of artificial neural networks, which often serve as equivalent models (and often considerably extending potential spectrum of applications) in relation to traditional statistical methods, play the key role here.