



Michał Palmąka

Zastosowanie sieci neuronowych w geologii

Application of artificial neural networks in geology

Streszczenie: Szybki rozwój technik badawczych w geologii oraz możliwości gromadzenia dużej liczby danych dotyczących środowiska gruntowo-wodnego wymagają korzystania z nowych metod ich analizy. W artykule przedstawiono podstawowe informacje dotyczące zasad funkcjonowania wybranych rodzajów sieci neuronowych oraz ich cechy, które umożliwią efektywne przetwarzanie licznych zbiorów danych. Wskazano na możliwość zastosowania modeli neuronowych do rozwiązywania problemów regresyjnych i klasyfikacyjnych oraz połączenie tego typu analiz z systemem informacji przestrzennej (GIS). Prezentowane metody poparto wybranymi przykładami literaturowymi dotyczącymi zastosowania sieci neuronowych w geologii.

Słowa kluczowe: symulacje nieliniowe, sieci neuronowe, perceptrony wielowarstwowe

Abstract: The rapid development of research techniques in geology and storage of a huge amount of digital data concerning a non-biotic environment require new methods to analyze and to draw correct conclusions from available information. This paper presents characteristics and basic principles of running and operating artificial neural networks (ANN), which enable to effectively process numerous data sets. Different geological applications of neural systems are discussed in the context of regression models and classification issues. Additionally some efficient neural network models with a Geographical Information System (GIS) combination are described and proved as a complex spatial methodology.

Key words: nonlinear simulations, neural network, multi-layer perceptron (MLP)

Wprowadzenie

Sieci neuronowe są obliczeniowymi strukturami opartymi na zasadzie funkcjonowania ludzkiego mózgu. Charakteryzują się one szybkim i efektywnym przetwarzaniem zbiorów danych o dużej liczebności. Przetwarzanie informacji ma charakter rozproszony i odbywa się równolegle. Cechą charakterystyczną sieci

Michał Palmąka, Uniwersytet Warszawski, Wydział Geologii, Katedra Ochrony Środowiska i Zasobów Naturalnych, ul. Żwirki i Wigury 93, 02-089 Warszawa, e-mail: mpalma-ka@uw.edu.pl

neuronowych jest możliwość generalizacji procesów nieliniowych, co istotnie rozszerza spektrum ich zastosowań. Odpowiednio zaprojektowane sieci neuronowe wykazują dużą tolerancję na błędy lub braki danych wejściowych, które służą do rozwiązania danego problemu.

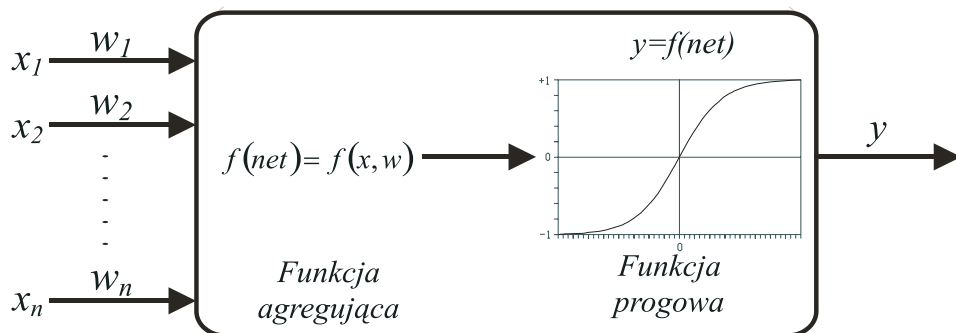
Badania nad sztucznymi sieciami neuronowymi rozpoczął w latach 40. XX w. zespół McCullocha (McCulloch i in., 1943), a w latach 50. były rozwijane przez Rosenblatta (Rosenblatt, 1962). Jednakże dopiero około roku 1985 zyskały na popularności dzięki wprowadzeniu i przetestowaniu przez zespół Rumelharta (Rumelhart i in., 1986) metody wstecznej propagacji błędu (back propagation) jako skutecznego algorytmu uczenia sieci neuronowej. Pierwotnie sieci neuronowe zastosowano w dziedzinach ekonomicznych, jednakże ich możliwości generalizacji procesów znalazły uznanie także w innych gałęziach nauki. Wraz z rozwojem sieci neuronowych (ich architektury i algorytmów uczących) obserwuje się wdrażanie urządzeń oraz systemów sterowania procesami technologicznymi, w których wykorzystuje się tego typu struktury obliczeniowe, do przemysłu (np. optymalizacja procesu spalania w elektrowniach, kontrola instalacji przesyłowych) i życia codziennego (np. zabezpieczenia biometryczne).

Budowa sztucznego neuronu

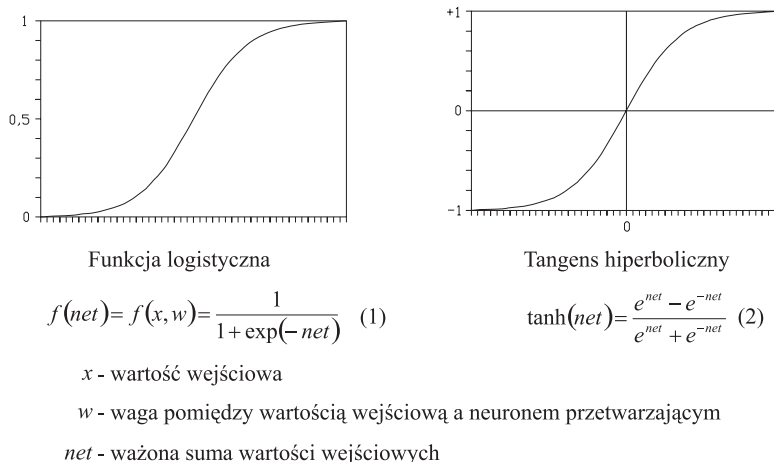
Sieć neuronowa zbudowana jest ze zbioru pojedynczych neuronów oraz połączeń między nimi. Neuron jest podstawowym elementem przetwarzającym, który posiada określoną liczbę wejść. W przypadku pojedynczego neuronu wejścia stanowią połączenia, którym przypisane są wagi (ryc. 1).

Wewnątrz neuronu przeprowadzane są dwie operacje:

- agregowanie wartości wejściowej do neuronu (funkcja potencjału postsynaptycznego);
- wyznaczenie wartości wyjściowej z neuronu (funkcja aktywacji; funkcja progowa).



Ryc. 1. Schemat budowy sztucznego neuronu
Fig. 1. Model of an artificial neuron structure



Ryc. 2. Typowe funkcje aktywacji wykorzystywane w sieciach neuronowych
 Fig. 2. Typical activation functions applied in neural networks

Do dwóch najczęściej stosowanych funkcji aktywacji należy zaliczyć: funkcję logistyczną i tangens hiperboliczny, których wykresy i wzory zostały przedstawione na rycinie 2.

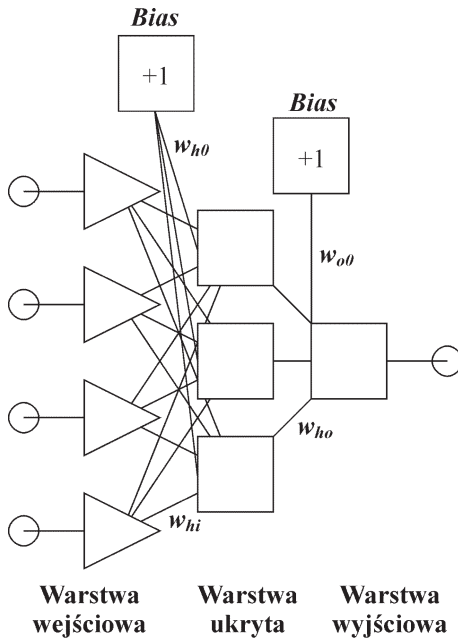
Zarówno sigmoidalna funkcja logistyczna, jak i tangens hiperboliczny mają charakter nieliniowy i są łatwo różniczkowalne, co ułatwia obliczenia w trakcie procesu uczenia i minimalizacji błędu sieci. Generowane przez funkcję logistyczną wartości wyjściowe są z przedziału $(0;1)$, natomiast przez \tanh z przedziału $(-1;1)$, dzięki temu w łatwy sposób można konwertować otrzymane wyniki do wartości rzeczywistych. Sporadycznie wykorzystuje się inne nieliniowe funkcje aktywacji, m.in.: funkcję wykładniczą, pierwiastek kwadratowy, sinus.

Topologia sieci

W większości zastosowań neurony zgrupowane są w warstwy (ryc. 3). Ze względu na wzajemne połączenia można wyróżnić trzy rodzaje warstw:

- wejściową, która reprezentuje parametry wejściowe (zmiennie niezależne) – wartości wyjściowe są przekazywane do neuronów drugiej warstwy bez przetworzenia;
- ukrytą, w której neurony posiadają pewną liczbę połączeń dochodzących do nich od neuronów warstwy poprzedniej, jak również łączą się z neuronami warstwy następczej;
- wyjściową, charakteryzującą poszukiwane wartości (zmiennie zależne).

Dane wejściowe z pierwszej warstwy są ważone i przesyłane do neuronów przetwarzających w następnej warstwie (ukrytej lub wyjściowej). W ramach przetwarzania informacji w każdym neuronie dochodzi do agregowania ważonych wartości, klasyfikacji zgodnie z funkcją progową neuronu i wysłania wartości wyjściowej



Ryc. 3. Schemat budowy sztucznej sieci neuronowej – perceptron wielowarstwowy (MLP)

Fig. 3. Model of an artificial neural network – multi-layer perceptron (MLP)

do wszystkich neuronów w kolejnej warstwie. Zatem wartość wyjściowa z j -tego neuronu przetwarzającego jest określona przez wyrażenie:

$$o_j = f\left(\sum w_{ji} x_i + w_{j0} \beta_j\right) \quad (3)$$

o_j – wartość wyjściowa,

f – funkcja progowa,

w_{ji} – waga pomiędzy i -tą wartością wejściową a j -tym neuronem przetwarzającym,

x_i – i -ta wartość wejściowa,

w_{j0} – waga pomiędzy neuronem składnika stałego a j -tym neuronem przetwarzającym,

β_j – neuron składnika stałego.

Proces uczenia sieci

Algorytmy uczące minimalizują błędy poprzez zmianę wartości wag pomiędzy neuronami. Zmiana wag i wyznaczanie błędu dokonywane jest iteracyjnie. Sposób uczenia sieci neuronowej (dopasowania wag połączeń) jest zależny od rodzaju danych oraz procedury uczącej. Ze względu na charakter zagadnienia i rodzaj danych można wyróżnić:

- uczenie nadzorowane (typowe dla sieci perceptronowych i rekurencyjnych),
- uczenie bez nadzoru (typowe dla sieci Kohonena).

Istnieje szereg algorytmów modyfikujących wagi połączeń pomiędzy neuronami. Proces uczenia sieci jest numeryczną procedurą minimalizacji funkcji celu, która w klasycznym ujęciu jest błędem średniokwadratowym, rozumianym jako różnica pomiędzy wartościami wyjść generowanymi przez sieć dla określonego wektora wejść a wartościami wzorcowymi. Jak już powyżej stwierdzono, błąd j -tego neuronu może być zdefiniowany jako kwadrat różnicy pomiędzy wartością generowaną przez sztuczną sieć neuronową a wartością wzorcową. Wyrażenie zgodnie z zaproponowaną definicją ma postać (Rumelhart i in., 1986):

$$E = \frac{1}{2} (d_j - o_j)^2 \quad (4)$$

E – błąd,

d_j – pożądana wartość j -tego neuronu,

o_j – wygenerowana wartość j -tego neuronu.

Wagi mogą być modyfikowane po obliczeniu wartości wyjściowej dla każdego przypadku w zbiorze uczącym (uczenie przyrostowe – *incremental learning*) lub po przeanalizowaniu całego zbioru (uczenie wsadowe – *batch learning*). Do najpopularniejszej metody modyfikacji wag połączeń pomiędzy sztucznymi neuronami należy algorytm wstecznej propagacji błędu (*back propagation*) z licznymi jego modyfikacjami. Algorytm wstecznej propagacji błędu jest uogólnieniem reguły delty i należy do metod gradientowych. W założeniu algorytm wykorzystuje prawidłowość stwierdzającą, że gradient funkcji wskazuje kierunek jej najszybszego wzrostu, a po zmianie znaku na minus – najszybszego spadku wartości funkcji. Właściwość ta pozwala na minimalizację funkcji celu przez modyfikację jej zmiennych (wag połączeń sieci), w kierunku najszybszego spadku wartości funkcji, czyli proporcjonalnie do gradientu funkcji celu. Spośród wielu innych algorytmów uczących można wymienić także metodę gradientów sprzężonych, metodę quasi-Newtona, metodę Levenberga-Marquardta, metodę szybkiej propagacji błędu, metodę Delta-Bar-Delta.

W celu kontroli powstających błędów oraz poprawności procesu uczenia (np. wykluczenie przeuczenia sieci polegającego na zaniku zdolności do generalizacji analizowanego procesu) cały zbiór danych wejściowych dzielony jest na trzy podzbiory: uczący, walidacyjny i testowy.

Zastosowanie sieci neuronowych

Najczęstszym zastosowaniem sieci neuronowych jest rozwiązywanie problemów regresyjnych lub klasyfikacyjnych. W zagadnieniach regresyjnych celem jest oszacowanie wartości ciągłej zmiennej wyjściowej w sytuacji, gdy znane są wartości zmiennych wejściowych. Z kolei w klasyfikacji celem stawianym sieci jest przypisanie każdego przypadku reprezentowanego przez odpowiedni zestaw danych wejściowych do jednej z wybranych klas – zadaniem sieci jest estymacja prawdopodobieństwa przynależności przypadku do danej klasy.

Przykładem zastosowania sztucznych sieci neuronowych jest model osiadania powierzchni terenu spowodowanego eksploatacją podziemną węgla kamiennego w Velenje (Słowenia) (Ambrožič, Turk, 2003). Z wykorzystaniem sieci neuronowych utworzono monitoring i dokonano oceny zagrożenia wystąpienia ruchów masowych w okolicy Okharpauwa wzdłuż autostrady Katmandu-Trishuli w Nepalu (Achet, Neaupane, 2004). W geotechnice sieci neuronowe znajdują zastosowanie do zaawansowanych aproksymacji charakterystyk odkształcenia gruntu (Turk i in., 2000). Z wykorzystaniem omawianego aparatu obliczeniowego możliwa jest ocena wytrzymałości gruntów na ścinanie (Attoh-Okine, Fekpe, 1996) oraz analiza wy-

trzymałości gruntów w obrębie jednej grupy utworów (np. laterytów), które charakteryzują się dużą niejednorodnością parametrów fizycznych i chemicznych. Modele sieci neuronowych pozwalają na opracowywanie i szybką interpretację badań polowych, np.: wyników sondowań statycznych CPT (Goh, 1995), prognozowanie osiadań gruntu spowodowanych tunnelingiem (Shi i in., 1998), a także na dokonywanie geoinżynierskich szacowań, takich jak obliczenia nośności pali (Lee, Lee, 1996) oraz ocena zagęszczenia nasypów (Sulewska, 2006). Sztuczne sieci neuronowe wykorzystywane są też do prognozowania wystąpienia wstrząsów podziemnych i tąpnięć w kopalniach (Rudajev, Čiž, 1999), określania współczynnika filtracji dla różnych typów utworów (Wong i in., 1999). W zagadnieniach geosrodowiskowych sztuczne sieci neuronowe są używane m.in. do prognozowania ilości wyłukiwanych metali ciężkich ze składowanych osadów ściekowych (Du i in., 1994). Powiązanie sztucznych sieci neuronowych i GIS (*Geographic Information System – GIS*) wykorzystano jako alternatywną metodę do interpolacji wartości stężeń radioaktywnego ^{137}Cs w rejonie Czarnobyli (Kanevski i in., 2004). Sieci neuronowe znalazły zastosowanie w prognozowaniu zmian jakości wód podziemnych w układzie przestrzennym oraz klasyfikacji punktów monitoringowych do obszaru o określonym użytkowaniu terenu w zlewni górnej Wisły (Kmieciak, 2004). Przetwarzanie danych za pomocą sieci neuronowych posłużyło do optymalizacji parametrów pomp ciepła w systemach płytkowej geotermii na terenie Polski (Tkacz, 2003). Przy wykorzystaniu rozproszonych informacji przestrzennych z lat 1976–2005 i perceptronów wielowarstwowych możliwe było utworzenie modelu osiadania terenu wokół KWB Bełchatów (Palmąka, 2007). Dodatkowo dzięki zastosowaniu klasyfikującej sieci Kohonena i nieliniowego skalowania wielowymiarowego (*Sammon's mapping*) dokonano analizy przestrzennej pod kątem występowania ewentualnych subregionów osiadań.

Jednym z przykładów wykorzystania perceptronu wielowarstwowego do klasyfikacji jest identyfikacja rodzajów gruntów i ich przyporządkowanie ze względu na przydatność do celów budowlanych (Cal, 1995). Na podstawie trzech cech gruntu: współczynnika plastyczności, wilgotności przy granicy płynności i zawartości minerałów ilastych, możliwa była poprawna klasyfikacja rodzajów gruntów. Ze względu na adaptacyjny charakter sztucznych sieci neuronowych oraz możliwość połączenia tej metody obliczeniowej z GIS możliwe stało się przeprowadzanie oceny ryzyka związanego z zagrożeniem osuwiskami w czasie rzeczywistym (Aleotti, Chowdhury, 1999). Skuteczność połączenia tych dwóch narzędzi analitycznych została sprawdzona w przypadku analizy ryzyka uruchomienia osuwisk w rejonie Boun (Korea) (Lee i in., 2003), jak również wywołania ruchów masowych spowodowanych wstrząsami sejsmicznymi w rejonie Baguio City (Filipiny) (Lee, Ewangelista, 2006). Sztuczne sieci neuronowe zostały zaadaptowane także do określania ryzyka związanego z powstawaniem nieciągłych uskoków (spękania i szczeliny na powierzchni terenu) w rejonie Yuci City, w Chinach (Wu i in., 2003).

Podsumowanie

Sztuczne sieci neuronowe stanowią silną i coraz szerzej stosowaną grupę metod prognostycznych i klasyfikacyjnych, która jest szczególnie przydatna w przypadkach, gdy brak jest dostatecznej wiedzy na temat niezależnych czynników wpływających na analizowany proces. Modele neuronowe mają charakter adaptacyjny, co oznacza, że mogą służyć do opisu zależności zmieniających się w czasie. W sytuacji pojawienia się nowych danych może być przeprowadzony proces douczenia sieci, co umożliwi uwzględnienie w tworzonym modelu informacji zawartych w najnowszych obserwacjach.

Istotny wpływ na otrzymywane wyniki ma odpowiednia topologia sieci oraz zastosowane metody w trakcie procesu uczenia. Należy pamiętać, że zoptymalizowane wartości wag sztucznej sieci neuronowej nie mogą być w prosty sposób interpretowane tak, jak ma to miejsce przy analizie klasycznych modeli regresyjnych. Z tego też powodu sieci neuronowe są bardziej przydatne w zagadnieniach prognostycznych niż w szczegółowych analizach procesu (model czarnej skrzynki). Sieć neuronowa jako model posiada wiele parametrów ograniczających (wagi połączeń, funkcje agregowania i aktywacji), w związku z tym wymagane jest w trakcie procesu uczenia dostarczenie dużej ilości informacji (obserwacji) na temat badanego zjawiska w celu uniknięcia przeuczenia sieci lub niedoboru stopni swobody w modelu. Ze względu na iteracyjny charakter przeprowadzanych obliczeń algorytmu uczącego, proces uczenia sieci jest czasochłonny.

Przytoczone przypadki oraz zauważalny wzrost wykorzystania sztucznych sieci neuronowych w różnych dziedzinach geologii świadczą o skuteczności tej struktury obliczeniowej w rozwiązywaniu problemów regresyjnych, klasyfikacyjnych oraz optymalizacyjnych i zachęcają do podejmowania prób ich zastosowania.

Literatura

- Achet S.H., Neaupane K.M., 2004. Use of backpropagation neural network for landslide monitoring: a case study in higher Himalaya. *Engineering Geology* 74: 213–226.
- Aleotti P., Chowdhury R., 1999. Landslide hazard assessment: summary review and new perspectives. *Bulletin of Engineering Geology and the Environment* 58: 21–44.
- Ambrožič T., Turk G., 2003. Prediction of subsidence due to underground mining by artificial neural networks. *Computers & Geosciences* 29: 627–637.
- Attoch-Okine N.O., Fekpe E.S.K., 1996. Strength characteristic modeling of lateritic soils using adaptive neural networks. *Construction and Building Materials* 10 (8): 577–582.
- Du Y.G., Sreekrishnan T.R., Campbell P.G.C., 1994. Effect of pH on metal solubilization from sewage sludge: A neural-net-based approach. *Canadian Journal of Civil Engineering* 21: 728–735.
- Goh A.T.C., 1995. Neural networks for evaluating CPT calibration chamber test data. *Microcomputers in Civil Engineering* 10 (2): 147–151.

- Kanevski M., Parkin R., Podnukhov A., Timonin V., Maignan M., Demyanov V., Canu S., 2004. Environmental data mining and modeling based on machine learning algorithms and geostatistics. *Environmental Modeling & Software* 19: 845–855.
- Kmieciak E., 2004: Prognozowanie zmian jakości wód podziemnych w układzie przestrzennym z wykorzystaniem sieci neuronowych. *Biuletyn Państwowego Instytutu Geologicznego* 412: 5–70.
- Lee S., Evangelista D.G., 2006. Earthquake-induced landslide-susceptibility mapping using an artificial neural network. *Natural Hazards and Earth System Sciences* 6: 687–695.
- Lee I.M., Lee J.H., 1996: Prediction of pile bearing capacity using artificial neural networks. *Computers and Geotechnics* 18 (3): 189–200.
- Lee S., Ryu J.H., Lee M.J., Won J.S., 2003. Use of an artificial neural network for analysis of the susceptibility to landslides at Boun, Korea. *Environmental Geology* 44: 820–833.
- McCulloch, Warren S., Pitts W., 1943. A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* 5: 115–133.
- Palmąka M., 2007. Modelowanie osiadania terenu wokół KWB Bełchatów z wykorzystaniem analiz wielowymiarowych i GIS. Uniwersytet Warszawski. Praca doktorska (w przygotowaniu).
- Rosenblatt D.E., 1962. *Principles of Neurodynamics*. Spartan Press, New York.
- Rudajev V., Číž R., 1999. Estimation of mining tremor occurrence by using neural networks. *Pure and Applied Geophysics* 154 (1): 57–72.
- Rumelhart D., McClelland J., 1986. *Parallel distributed processing Explorations in the Microstructure of Cognition*. Vol. 1: Foundation. MIT Press, Cambridge, Mass.
- Shi J., Ortiago J.A.R., Bai J., 1998. Modular neural network for predicting settlements during tunneling. *Journal of Geotechnical and Geoenvironmental Engineering, ASCE* 124 (5): 389–394.
- Sulewska M.J., 2006. O zastosowaniu sztucznych sieci neuronowych w ocenie zagęszczania nasypów. *Inżynieria i Budownictwo* 6: 337–338.
- Tkacz M., 2003. Modelowanie warunków georodowiskowych przy pomocy metod sztucznej inteligencji na przykładzie hybrydowego systemu płytkiej geotermiki. Katedra Geologii Podstawowej, Wydział Nauk o Ziemi, Uniwersytet Śląski, Katowice. Rozprawa doktorska.
- Turk G., Logar J., Majes B., 2001. Modelling soil behaviour in uniaxial strain conditions by neural networks. *Advances in Engineering Software* 32 (10-11): 57–64.
- Wong P.M., Jang M., Cho S., Gedeon T.D., 2000. Multiple permeability predictions using an observational learning algorithm. *Computers and Geosciences* 26: 907–913.
- Wu Q., Ye S., Wu X., Chen P., 2003. A nonlinear modeling and forecasting system of earth fractures based on coupling of artificial neural network and geographical information system – exemplified by earth fractures in Yuci City, Shanxi. *Environmental Geology* 45: 124–131.