

Karolina Bartos

Pozyskiwanie wiedzy z danych za pomocą sieci neuronowych

Ekonomiczne Problemy Usług nr 68, 339-346

2011

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

KAROLINA BARTOS

Uniwersytet Ekonomiczny we Wrocławiu

POZYSKIWANIE WIEDZY Z DANYCH ZA POMOCĄ SIECI NEURONOWYCH

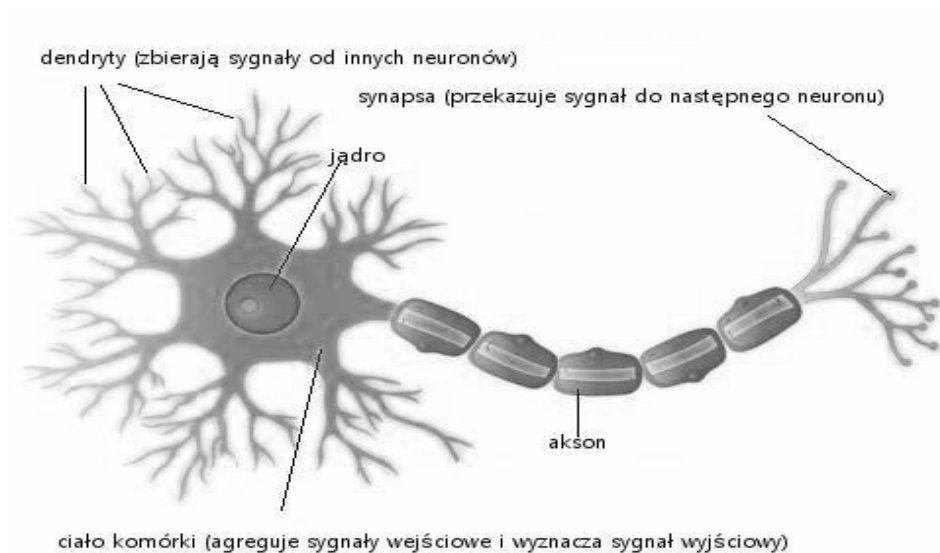
Wprowadzenie

Ogólnie wiadomo, że wiedza jest źródłem przewagi konkurencyjnej przedsiębiorstwa na rynku. Dlatego, próbując ją pozyskać, tworzy się ogromne bazy zawierające coraz więcej różnego typu danych. Ze względu na ich ilość tradycyjne techniki analizy stają się zawodne. Pojawiają się więc nowe narzędzia służące do wydobywania z morza danych istotnych informacji. Jednym z nich są sztuczne sieci neuronowe (*artificial neural networks*), zwane w skrócie także sieciami neuronowymi (*neural networks*). Należą one do zaawansowanej techniki modelowania, zdolnej do odwzorowywania nadzwyczaj złożonych funkcji, a ich nieliniowy charakter istotnie wzbogaca możliwości zastosowania.

1. Geneza sztucznych sieci neuronowych

Inspiracją do stworzenia sztucznych sieci neuronowych była wiedza z dziedziny nauk biologicznych na temat budowy i sposobu funkcjonowania systemu uczącego się w mózgach zwierząt. Składa się on z połączonych ze sobą neuronów tworzących sieć. Neurony przekazują sobie informacje za pomocą impulsu elektrochemicznego. Wychwytyją go od sąsiednich neuronów za pomocą dendrytów, następnie ciało komórki (*perikarion*) agreguje sygnały wejściowe i wyznacza sygnał wyjściowy, który aksonem zostaje wyprowadzony do synaps. Te przekazują go dalej, za pomocą skomplikowanych procesów elektryczno-chemicznych, do kolejnych neuronów. Neurony tworzą sieć, wykorzystując połączenia synaptyczne, które

łączą zakończenie aksonu jednego neuronu z dendrytem następnego. Na rysunku 1 przedstawiono schemat i funkcjonowanie biologicznego neuronu.



Rys. 1. Schemat budowy biologicznego neuronu i jego funkcjonowanie

Źródło: opracowanie własne na podstawie: R. Tadeusiewicz: *Wprowadzenie do praktyk stosowania sieci neuronowych*, w: *Sieci neuronowe*. Materiały konferencyjne, StatSoft Polska, Kraków 1999, s. 11–12.

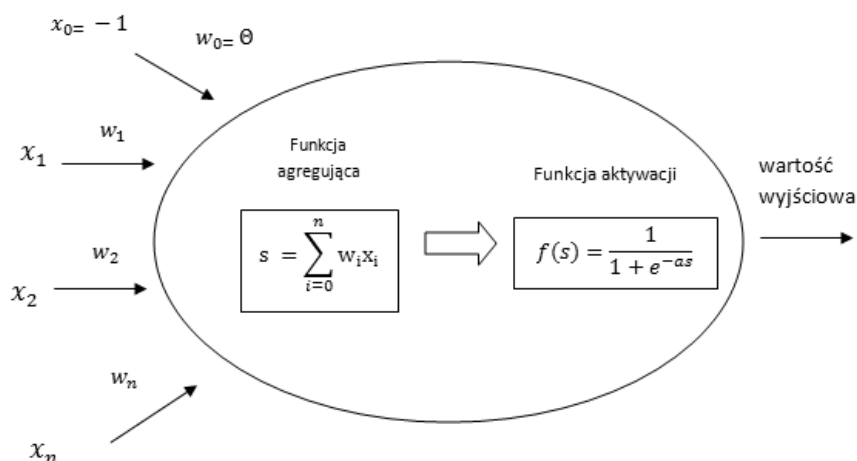
Mózg człowieka zawiera około 10^{11} neuronów, każdy połączony średnio z 10 tys. innych neuronów, co daje łącznie 10^{15} połączeń synaptycznych¹. Mimo że pojedynczy neuron ma prostą budowę, to gęsta siatka połączonych neuronów jest w stanie wykonać skomplikowane zadania uczenia się.

Podobną budowę mają sztuczne sieci neuronowe, które także składają się z ze wzajemnie połączonych ze sobą elementów przetwarzających informacje. Jednak są one jedynie modelem matematycznym naśladującym fragmenty rzeczywistego mózgu w bardzo uproszczony sposób.

¹ D.T. Larose: *Odkrywanie wiedzy z danych. Wprowadzenie do eksploracji danych*, Wydawnictwo Naukowe PWN, Warszawa 2006, s. 131.

2. Model sztucznego neuronu i sposób jego funkcjonowania

Sztuczny neuron jest prostym systemem przetwarzającym wartości sygnałów wprowadzonych na jego wejściu w pojedynczą wartość wyjściową. Jest więc układem posiadającym, zależnie od liczby danych wejściowych, n wejść (podobnie jak n dendrytów w neuronie biologicznym) i zawsze dokładnie jedno wyjście (akson). Model sztucznego neuronu przedstawiony jest na rysunku 2.



Rys. 2. Schemat budowy sztucznego neuronu z przykładową funkcją agregującą i aktywacji oraz z uwzględnieniem wartości progowej jako wagi w_0

Źródło: opracowanie własne.

Sygnaly wejściowe (oznaczone jako: x_1, x_2, \dots, x_n) są wartościami danych pierwotnych, podawanych do sieci z zewnątrz lub wartościami pośrednimi, pochodzącymi z połączonych z danym neuronem wyjść innych neuronów. Wartości te są przemnażane przez odpowiednie współczynniki zwane wagami² (oznaczone jako: w_1, w_2, \dots, w_n), które podobnie jak w neuronie biologicznym, odzwierciedlają siłę reakcji synapsy. „Ważone” sygnały wejściowe są sumowane i za pomocą funkcji aktywacji przetwarzane w wartość wyjściową (sygnał wyjściowy). W najprostszym przypadku funkcja aktywacji przyjmuje postać funkcji identycznościowej: $f(s) = s$,

² Proces ten może się różnić w zależności od rodzaju funkcji agregującej. Na rysunku 2 pokazana jest najpopularniejsza funkcja agregująca.

która w połączeniu z przedstawioną na rysunku 2 funkcją agregacji definiuje funkcjonowanie tzw. neuronu liniowego, pozwalającego na budowę bardzo użytecznych liniowych sieci neuronowych³. Innym przykładem funkcji aktywacji jest bardzo często stosowana funkcja sigmoidalna (rysunek 2), a także funkcja tangensoidalna i tzw. funkcje radialne (gaussowskie).

Każdy neuron posiada tzw. wartość progową (próg = Θ), określającą, jak silnie musi być jego pobudzenie, by doszło do uaktywnienia neuronu. Próg odejmowany jest od ważonej sumy wartości wejściowych:

$$s = \sum_{i=1}^k w_i x_i - \Theta$$

Neuron, dla którego $s \geq 0$, nazywamy pobudzonym. W procesie *uczenia się*⁴ sieci następuje modyfikacja parametrów poszczególnych neuronów, tj. wag (w_i) oraz wartości progowej (Θ). Wyodrębnienie Θ jako osobnego parametru wymagałoby prowadzenia dla niego osobnej notacji. Dlatego dogodniejsze jest włączenie progu do sumy ważonej poprzez stworzenie dodatkowej stałej wartości wejściowej $x_0 = -1$ oraz odpowiadającej jej wadze $w_0 = \Theta$. Dzięki włączeniu wartości progowej do wektora wag, podczas procesu uczenia próg będzie traktowany jak każdy inny parametr⁵. Jednocześnie ułatwi to zapis funkcji agregującej, która przyjmie postać jak na rysunku 2.

3. Budowa i działanie sztucznych sieci neuronowych

Możliwości przetwarzania informacji za pomocą pojedynczego neuronu są niewielkie, dlatego łączy się je ze sobą w grupy tworzące tzw. sztuczną sieć neuronową. Najczęściej stosowane są sieci należące do klasy jednokierunkowych sieci wielowarstwowych. Schemat budowy sieci tego typu jest pokazany na rysunku 3. Neurony uporządkowane są tam w warstwach: warstwa wejściowa, warstwy (warstwa) ukryte oraz warstwa wyjściowa. Sygnały przenoszące informację przepływają w jednym kierunku⁶ od warstwy wejściowej, przez kolejne warstwy ukryte do warstwy wyjściowej. Neurony warstwy wejściowej otrzymują na wejściu dane z ze-

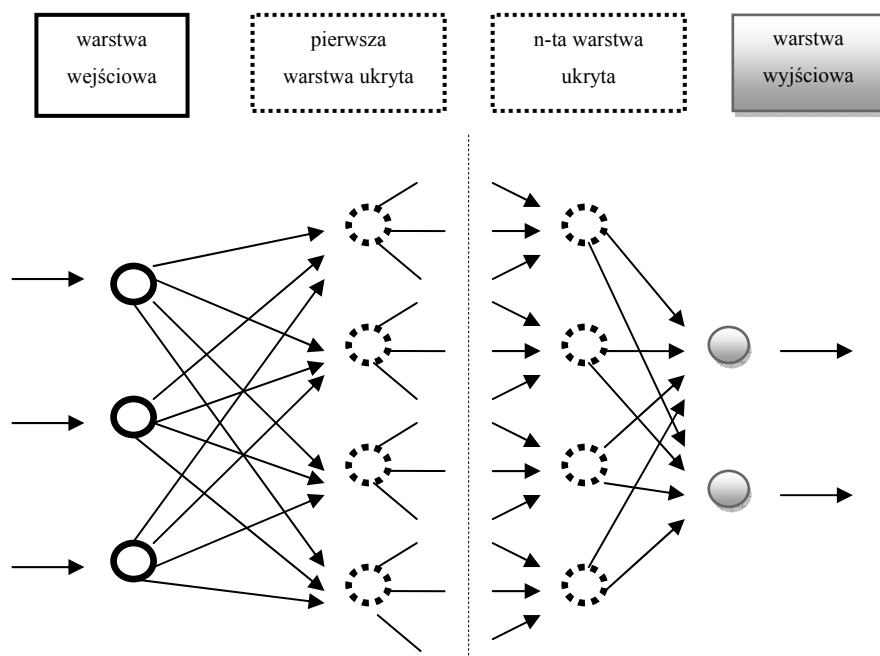
³ P. Lula, G. Paliwoda-Pękosz, R. Tadeusiewicz: *Metody sztucznej inteligencji i ich zastosowania w ekonomii i zarządzaniu*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2007, s. 80.

⁴ Proces uczenia zostanie omówiony w dalszej części artykułu.

⁵ K. Krawiec, J. Stefanowski: *Uczenie maszynowe i sieci neuronowe*, Wydawnictwo Politechniki Poznańskiej, Poznań 2003, s. 84–85.

⁶ Wśród sieci warstwowych istnieją również sieci rekurencyjne, w których wartości wyjściowe neuronów jednej warstwy są przekazywane na wejścia warstw wcześniejszych, a więc istnieją tzw. sprzężenia zwrotne. Przykładem sieci, w której wszystkie połączenia mają charakter sprzężeń zwrotnych, jest sieć Hopfielda.

wewnątrz sieci i przetwarzając je za pomocą funkcji agregującej i aktywacji, przekazują sygnały na wejścia neuronów pierwszej warstwy ukrytej. Te, dokonując kolejnych przekształceń, wysyłają przetworzone sygnały dalej – na kolejne wejścia następnych ukrytych warstw. W końcu sygnały docierają do warstwy wyjściowej. Tam, po przetworzeniu przez neurony budujące tę warstwę, wysyłane są na zewnątrz, jako już końcowe wartości wyjściowe sieci.



Rys. 3. Przykładowa struktura jednokierunkowej sieci wielowarstwowej

Źródło: opracowanie własne.

W sieciach warstwowych zwykle każdy neuron danej warstwy jest połączony ze wszystkimi neuronami warstwy bezpośrednio sąsiadującej. Taki typ połączeń jest nazywany *typem połączeń każdy-z-każdym*⁷.

Proces przygotowania sieci do jej prawidłowego działania nazywa się uczeniem. Polega on na modyfikacji przez sieć wartości wag (na rysunku 2 oznaczonych jako: w_0, w_1, \dots, w_n) na podstawie analizy danych ze *zbioru uczącego*. Sieć, opierając się na przedstawionych jej rzeczywistych przypadkach (wartościach danych ze zbioru uczącego), próbuje odkryć prawidłowości charakteryzujące te obiekty lub

⁷ P. Lula, G. Paliwoda-Pękosz, R. Tadeusiewicz: *Metody sztucznej...*, op. cit., s. 83–84.

badane zjawiska i zapamiętuje je w postaci zmodyfikowanych wartości wag⁸. Jeśli podczas uczenia wartość współczynnika wagowego przyjmie wartość zero, to dane połączenie pomiędzy neuronami zostanie usunięte. Jeśli zaś wszystkie wejściowe i wyjściowe wagi połączeń danego neuronu będą równe zero, to cały neuron, jako niepotrzebny zostaje usunięty. Zwykle jednokrotna prezentacja danych (jedna epoka) ze zbioru uczącego nie wystarcza do prawidłowego funkcjonowania sieci. Dlatego prezentowane są one wielokrotnie, często nawet kilka tysięcy lub milionów razy.

4. Przykładowe obszary zastosowań sztucznych sieci neuronowych do pozyskiwania wiedzy

Sieci neuronowe są narzędziem o wysokich możliwościach aplikacyjnych. Stosowane są do rozwiązywania różnego typu problemów, w wielu dziedzinach, takich jak np.: medycyna, zastosowania inżynierskie, geologia czy fizyka. W ekonomii znajdują szerokie zastosowanie, umożliwiając „odkrycie” z posiadanych danych niezbędnej wiedzy potrzebnej do podejmowania kluczowych decyzji. Do przykładowych klas problemów ekonomicznych, które z powodzeniem są rozwiązywane za pomocą sieci neuronowych, można zaliczyć: przewidywanie (predykcję), klasyfikację (przydział do istniejących klas) oraz grupowanie⁹.

4.1. Predykcja

Sieci często wykorzystywane są do przewidywania na podstawie danych wejściowych, jakie będą dane wyjściowe. Przykładami tego typu zastosowań mogą być np.: przewidywanie zachowań rynków finansowych (możliwości zmian lub kontynuacji trendu cenowego poszczególnych instrumentów, ale także ich konkretne ceny), prognozy ekonomiczne rozwoju przedsiębiorstw, przewidywanie zapotrzebowania na konkretny produkt lub usługi, a także prognoza zjawisk i wskaźników makroekonomicznych, takich jak np. bezrobocie czy inflacja. Do celów predykcyjnych zwykle wykorzystuje się najprostsze warstwowe sieci perceptronowe o jednokierunkowym przepływie sygnału, z nieliniowymi funkcjami aktywacji. Do bardziej zaawansowanych typów sieci używanych do tego typu problemów należą sieci stosujące regresję uogólnioną (GRNN).

⁸ *Ibidem*, s. 84–85.

⁹ Poniższe grupy problemów opisane na podstawie: R. Tadeusiewicz: *Sieci neuronowe*, Akademicka Oficyna Wydawnicza RM, Warszawa 1993, s. 16–18; J. S. Zieliński: *Inteligentne systemy w zarządzaniu. Teoria i praktyka*, Wydawnictwo Naukowe PWN, Warszawa 2000, s. 171–208.

4.2. Klasyfikacja (przydział do istniejących klas)

Polega na wytworzeniu przez sieci neuronowe podczas procesu uczenia pól decyzyjnych odpowiadających zdefiniowanym przez użytkownika klasom. Prawidłowo nauczona sieć potrafi na podstawie danych wejściowych poprawnie zakwalifikować dany obiekt do odpowiedniej klasy. Przykładem takiej klasyfikacji może być ocena zdolności kredytowej klienta. Sieć jest uczona na podstawie danych dotyczących wcześniejszych klientów, a następnie na podstawie danych zawartych we wniosku kredytowym obecnego klienta przydziela go do odpowiedniej klasy, np.: klient wypłacalny, niewypłacalny, pozostali. Do rozwiązywania zagadnień klasyfikacyjnych wykorzystuje się zwykle jednokierunkowe sieci z ciągłymi, nieliniowymi funkcjami aktywacji. Identyfikacja klas jest w nich zorganizowana najczęściej poprzez przypisanie jednemu neuronowi wyjściowemu sieci jednej konkretnej klasy. Sieci tego typu mają więc tyle neuronów wyjściowych, na ile klas został podzielony rozpatrywany problem.

4.3. Grupowanie

Polega na grupowaniu rekordów, obserwacji lub przypadków w klasy podobnych obiektów. Różni się od klasyfikowania tym, że tutaj nie znamy początkowo klas, do których będą przydzielane obiekty. Dopiero algorytm grupowania próbuje podzielić cały zbiór danych na klasy, w których podobieństwo obiektów wewnątrz klasy jest możliwie jak największe, a podobieństwo do obiektów spoza danej klasy jak najmniejsze. Przykładem zastosowania grupowania może być segmentacja klientów. Dzięki segmentacji można otrzymać informacje, jakie cechy posiadają klienci z danej grupy oraz czego oczekują od kupowanych przez nich produktów (usług). Pozwala to na lepsze dostosowanie oferty i strategii marketingowej do typu klientów danej grupy. Do rozwiązania problemu grupowania stosuje się różne typy sieci samouczących, najczęściej są to sieci Kohonena.

Podsumowanie

Nie sposób jest wymienić wszystkie możliwości zastosowań sieci do pozyskiwania wiedzy z danych, jednak z całą pewnością można stwierdzić, że jest to narzędzie bardzo chętnie stosowane. Wynika to nie tylko z jego dużej użyteczności, ale także z faktu, że sieci neuronowe automatycznie uczą się na podanych przez użytkownika przykładach i w praktyce same konstruują potrzebne modele. Jednak należy zauważyć, że niezbędna jest tutaj pewna wiedza użytkownika dotycząca odpowiedniego wyboru typu sieci, jej architektury (sposobu ułożenia neuronów, określenie ich liczby i rodzaju połączeń pomiędzy nimi), określenia funkcji agregującej i aktywacji w poszczególnych warstwach, ustalenie czasu uczenia się sieci (liczby epok), a także właściwej interpretacji uzyskanych wyników.

Literatura

1. Krawiec K., Stefanowski J.: *Uczenie maszynowe i sieci neuronowe*, Wydawnictwo Politechniki Poznańskiej, Poznań 2003.
2. Larose D.T.: *Odkrywanie wiedzy z danych. Wprowadzenie do eksploracji danych*, Wydawnictwo Naukowe PWN, Warszawa 2006.
3. Lula P., Paliwoda-Pękosz G., Tadeusiewicz R.: *Metody sztucznej inteligencji i ich zastosowania w ekonomii i zarządzaniu*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków 2007.
4. Tadeusiewicz R.: *Sieci neuronowe*, Akademicka Oficyna Wydawnicza RM, Warszawa 1993.
5. Tadeusiewicz R.: *Wprowadzenie do praktyk stosowania sieci neuronowych*, w: *Sieci neuronowe*, materiały konferencyjne, StatSoft Polska, Kraków 1999.
6. Zieliński J. S.: *Inteligentne systemy w zarządzaniu. Teoria i praktyka*, Wydawnictwo Naukowe PWN, Warszawa 2000.

KNOWLEDGE ACQUISITION FROM DATA USING NEURAL NETWORKS**Summary**

The aim of the following article is to present an artificial neural network as a tool for extracting useful knowledge from data. The paper includes basic information about neural networks: origin, structure and functioning. What is more, the author describes use of this tool to solve economic problems in three areas: prediction, classification and clustering.

Translated by Karolina Bartos