

Sztuczne sieci neuronowe Kohonena jako narzędzie w taksonomii paleontologicznej — metodyka oraz zastosowanie na przykładzie późnokredowych belemnitów

Zbigniew Remin¹



Artificial Kohonen neural networks as a tool in paleontological taxonomy — an introduction and application to Late Cretaceous belemnites. *Prz. Geol.*, 56: 58–66.

A b s t r a c t. Artificial neural networks (ANNs), the computer software or systems that are able to “learn” on the basis of previously collected input data sets are proposed here as a new useful tool in paleontological modeling. Initially ANNs were designed to imitate the structure and function of natural neural systems such as the human brain. They are commonly used in many natural researches such as physics, geophysics, chemistry, biology, applied ecology etc. Special emphasis is put on the Kohonen self-organizing mapping algorithm, used in unsupervised networks for ordination purposes. The application of ANNs for paleontology is exemplified by study of Late Cretaceous belemnites. The Kohonen networks objectively subdivided the belemnite material (~750 specimens)

into consistent groups that could be treated as monospecific. The possibility of transferring these results to the language of classical statistics is also presented. Further development and possibility of use of ANNs in various areas of paleontology, paleobiology and paleoecology is briefly discussed.

Keywords: paleontology, artificial intelligence, artificial neural networks, Kohonen neural networks, self-organizing map, classification, belemnites, Upper Cretaceous

Zagadnienia związane z obiektywną klasyfikacją taksonomiczną materiału kopalnego od zawsze stanowiły bardzo poważny problem. W znakomitej większości przypadków szczegółowe rozpoznanie danej formy kopalnej, szczególnie na niższych szczeblach taksonomicznych, bardzo często opiera się na naszym doświadczeniu oraz subiektywnych przesłankach. Dlatego takie rozpoznanie obarczone jest mniejszym lub większym błędem wynikającym z nieobiektywnego traktowania cech istotnych taksonomicznie. Problem ten w mniejszym stopniu dotyczy form kopalnych charakteryzujących się wyraźną i mało zmienną morfologią w obrębie np. gatunku. W takich przypadkach rozpoznanie nawet do szczebla gatunkowego nie następuje zwykle większych trudności. Problemy zaczynają się jednak, gdy musimy rozpoznawać formy o małej dyspersji morfologicznej (np. belemnity) lub charakteryzujące się dużą zmiennością wewnątrzrodzajową lub wewnątrzgatunkową.

Prawidłowe rozpoznanie i zaklasyfikowanie danego okazu ma daleko idące konsekwencje w dalszym wnioskowaniu. Po pierwsze, przybliża nas do rzeczywistego poznania zmienności, np. wewnątrzgatunkowej, a to z kolei przekłada się na możliwość śledzenia ewentualnych zmian filetycznych lub filogenetycznych w rozpatrywanych liniach ewolucyjnych. Po drugie, ma podstawowe znaczenie we wnioskowaniu paleoekologicznym i paleogeograficznym. Po trzecie, daje możliwość tworzenia wiarygodnych schematów biostratygraficznych.

Wymienione problemy zmuszają nas do poszukiwania nowych, bardziej obiektywnych narzędzi klasyfikacyjnych. Jednym z takich narzędzi są z pewnością sztuczne sieci neuronowe. W niniejszym artykule przedstawiono możliwości zastosowania sztucznych sieci neuronowych (ANN = *artificial neural networks*), w szczególności samoorganizujących się sieci Kohonena (SOM = *self-organizing maps*), do rozwiązywania problemów natury klasyfikacyjnej w badaniach organizmów kopalnych.

Przykładem zastosowania nowej metodyki jest studium taksonomiczne późnokredowych belemnitów z rodzajów *Belemnitella* d’Orbigny, 1840 oraz *Belemnella* Nowak, 1913 (Remin, 2007). W artykule zostały obszernie przedstawione podstawowe wiadomości i pojęcia na temat sztucznych sieci neuronowych, ze szczególnym uwzględnieniem samoorganizujących się sieci Kohonena. zilustrowano przykładowe sposoby analizy otrzymanych wyników oraz możliwość przełożenia ich na język klasycznej statystyki. Przeprowadzone badania oraz uzyskane rezultaty, których drobną część zaprezentowano w niniejszym artykule jako przykłady, stanowią część rozprawy doktorskiej autora pt: *Analiza paleontologiczna i znaczenie stratygraficzne belemnitów górnego kampanu i dolnego mastrychtu profilu doliny środkowej Wisły*.

Zastosowania sztucznych sieci neuronowych

Od lat 80. XX wieku ANN przeżywają gwałtowny rozwój i są wykorzystywane do rozwiązywania różnorodnych problemów z prawie wszystkich dziedzin nauki (medycyny, fizyki, chemii, biologii molekularnej, ekonomii, inżynierii, ekologii stosowanej itd.), głównie jednak problemów natury klasyfikacyjnej i predykcyjnej, w których dają bardzo dobre rezultaty. Dzieje się tak za sprawą zdolności ANN do odwzorowywania skomplikowanych, nieliniowych zależności między badanymi zjawiskami. Analiza takich zależności z zastosowaniem klasycznych metod statystycznych opartych na modelach liniowych powoduje wiele problemów natury metodologicznej.

Od połowy lat 90. XX wieku ANN są powszechnie używane do rozwiązywania złożonych problemów środowiskowych i ekologicznych. Przykłady zastosowań znaleźć można w różnych obszarach ekologii stosowanej, takich jak modelowanie efektu cieplarnianego (Seginer i in., 1994), modelowanie zmienności populacji ryb (Giske i in., 1998), przewidywanie produkcji fitoplanktonu (Scardi, 1996; Recknagel i in., 1997), przewidywanie różnorodności gatunkowej ryb (Guégan i in., 1998), klasyfikacja i zmienność populacji bezkręgowców (Chon i in., 1996,

¹Wydział Geologii, Uniwersytet Warszawski, ul. Żwirki i Wigury 93, 02-089 Warszawa; zbyh@uw.edu.pl

2000) lub wpływ zanieczyszczeń i regulacji rzek na rozmieszczenie i liczebność ryb (Kruk, 2007).

Zastosowania ANN w naukach o Ziemi są stosunkowo nieliczne, wydaje się jednak, że i tu w najbliższych latach będziemy obserwować wzrost zainteresowania wykorzystaniem ANN. Dotychczas ANN były stosowane np. do analizy profili wiertniczych (Baldwin i in., 1989; Rogers i in., 1992), korelacji poziomów pyłów wulkanicznych (Malmgren & Nordlund, 1996), w badaniach paleoceanograficznych (Malmgren & Nordlund, 1997), do klasyfikacji skał na podstawie danych geochemicznych (Kaminskas & Malmgren, 2004) i do identyfikacji zespołów litologicznych z danych wiertniczych (Chang i in., 2002). W badaniach paleontologicznych sztuczne sieci neuronowe praktycznie nie były dotychczas stosowane, a literatura ogranicza się do nielicznych, sympozjalnych doniesień; np. Marmo i in. (2006) stosując ANN uzyskali doskonałe rezultaty w automatycznym rozpoznawaniu otwornic. Większość z wymienionych prac świadczy o tym, że zastosowanie ANN daje lepsze rezultaty od klasycznych metod modelowania.

Czym są sztuczne sieci neuronowe?

Są to systemy komputerowe mające zdolność „uczenia się”. Jako aplikacje komputerowe ANN powstały na gruncie badań nad sztuczną inteligencją. Zgromadzone dane reprezentujące badany problem stanowią podstawę procesu uczenia z zastosowaniem określonego algorytmu uczenia. Ponieważ nie musimy samodzielnie programować wykorzystywanych ANN (możemy opierać się na wielu dostępnych aplikacjach komputerowych), wymagana jest od nas głównie wiedza empiryczna dotycząca badanych zjawisk i problemów, sposobu doboru oraz przygotowania danych, które następnie będą stanowiły przykłady dla uczącej się sieci neuronowej. Oczywiście, żeby właściwie wybrać sieć neuronową i należy zinterpretować otrzymane rezultaty, ważny jest dobór i znajomość odpowiednich typów algorytmów uczenia.

W ogólnych założeniach ANN naśladują pracę i działanie naturalnych systemów nerwowych (np. mózgu). Dlatego charakteryzują się kilkoma bardzo ważnymi cechami przynależnymi dotychczas tylko tym systemom:

1. **Uczą się na przykładach;** nabywają doświadczenia jak reagować na bodźce (dane wejściowe), które dostarczane są do mózgu (sieci neuronowej), powoduje to konkretne zachowania (dane wyjściowe).
2. Wyuczone sieci neuronowe (mózg lub sztuczna sieć neuronowa) są bardzo **odporne na niekompletność dostarczanych danych**. Oznacza to, że chociaż dysponujemy ograniczoną ilością danych, to możemy uzyskiwać zadowalające rezultaty. Przykładem znaczenia tej cechy może być nasza znajomość np. samochodów. Samochód możemy opisać wieloma różnorodnymi cechami, które będą charakteryzowały konkretny model. Często jednak wystarczy nam pojedynczy element nadwozia, aby zaklasyfikować obserwowane auto do konkretnej marki. Oczywiście, im większą ilość danych będziemy dysponować, tym precyzyjniejsza będzie klasyfikacja.

3. Sieci neuronowe mają **zdolność do generalizacji wiedzy**; na podstawie już zdobytych umiejętności w prawidłowy sposób reagują na nowe dane wejściowe (bodźce), pomimo że takie dane nie uczestniczyły nigdy w procesie uczenia. Cecha ta ma podstawowe znaczenie w uczących się systemach nerwowych. Przykładem wagi tej cechy jest nasza reakcja na spadający w kamieniołomie głaz. Choć nigdy wcześniej nie znaleźliśmy się w podobnej sytuacji i nigdy nie stwierdziliśmy empirycznie, jaki będzie efekt spotkania lecącego kamienia z naszą głową, to decyzje, które w naszym mózgu (sieci neuronowej) zapadną i spowodują reakcję naszego ciała na zaistniałą sytuację (czyli na dane wejściowe), będą słuszne — uskoczymy w bok (dane wyjściowe = efekt działania sieci). Na tym przykładzie widzimy, że wyuczone ANN potrafią klasyfikować sytuacje bądź obiekty (dane wejściowe), z którymi wcześniej nie miały do czynienia, a klasyfikacje te (dane wyjściowe) są prawidłowe.

Sztuczne sieci są zbudowane z **neuronów** — podstawowych jednostek przetwarzających dane — ułożonych w warstwy: 1) wejściową, której są prezentowane dane uczące; 2) wyjściową, na której są generowane wyniki, oraz ewentualnie 3) położonych między nimi warstw ukrytych. Każdy neuron przesyła sygnały do każdego neuronu warstwy następnej. Uczenie ANN jest iteracyjne. Jednostką czasu jest tzw. **epoka**. W trakcie poszczególnych epok sieci neuronowej jest prezentowany zbiór uczący, a rezultaty służą do modyfikacji wag i wartości progowych poszczególnych neuronów biorących udział w procesie uczenia. W ANN określany jest również współczynnik uczenia, który może być stały lub zmienny w czasie (w trakcie kolejnych epok).

Poszczególne rodzaje ANN (np. liniowe sieci neuronowe, perceptrony wielowarstwowe, sieci o radialnych funkcjach bazowych, sieci realizujące regresję uogólnioną, probabilistyczne sieci neuronowe oraz sieci Kohonena) różnią się między sobą, a decyzja, którą sieć zastosować, zależy głównie od rodzaju rozwiązywanego problemu badawczego. ANN uczy się, żeby minimalizować błąd sieci. Osiągane jest to przez rozłożoną w czasie (epoki) modyfikację parametrów sieci za pomocą określonych algorytmów (wstecznej propagacji błędów, gradientów sprzężonych, Quasi-Newtona, Lavenberga-Marquardta, szybkiej propagacji i innych). Najpopularniejszym algorytmem modyfikującym parametry sieci jest algorytm wstecznej propagacji błędów. Wyczerpujące informacje oraz dyskusje na temat uczenia neuronów, stosowania właściwych algorytmów i wyboru odpowiednich sieci można znaleźć w opracowaniach Pattersona (1996), Fausetta (1994) oraz np. w materiałach StatSoft (2001, 2006).

Większość ANN jest projektowana do nauki nadzorowanej w tzw. trybie z nauczycielem. Oznacza to, że uczenie polega na jednoczesnej prezentacji zbioru danych zawierających zmienne niezależne (wejściowych, np. dane biometryczne) oraz zmienne zależne (wyjściowe, np. konkretne gatunki), czyli przykłady i odpowiedzi. Znane są zatem wzorce, a trenowana sieć wychwytuje zależności między poszczególnymi zmiennymi. Prezentując następnie zbiór danych nieuczestniczących w procesie uczenia, sprawdzamy, czy sieci się to udało i czy w efekcie uzyskujemy prawidłowe odpowiedzi.

Samoorganizujące się sieci Kohonena

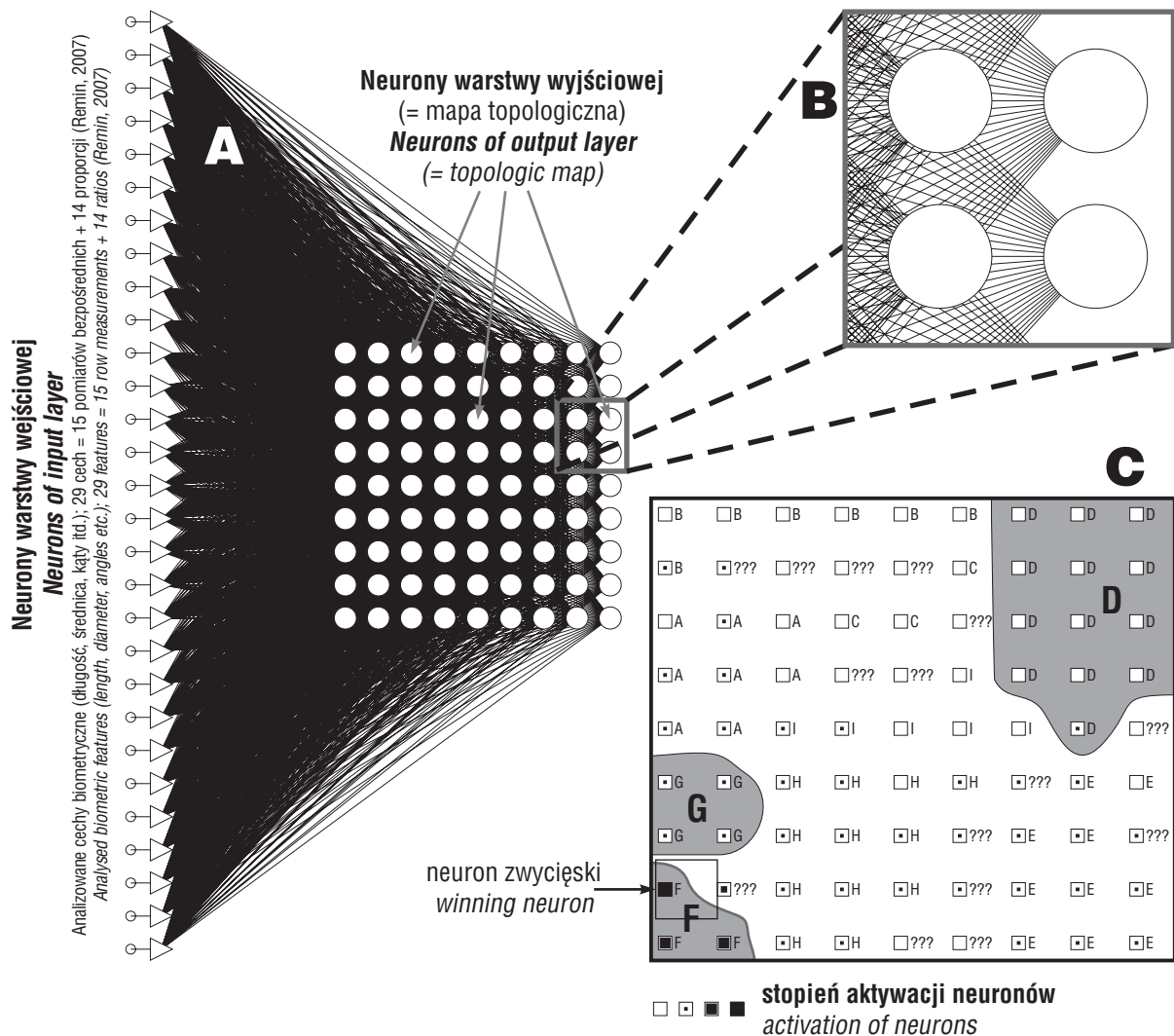
Sieci Kohonena wykorzystane przez autora do klasyfikacji belemnitów różnią się od innych sieci neuronowych architekturą i sposobem uczenia. Sieci Kohonena zaprojektowano do nauki nienadzorowanej w tzw. trybie bez nauczyciela (Kohonen, 1982; Fausett, 1994; Haykin, 1994; Patterson, 1996; materiały StatSoft 2001, 2006).

W sieciach projektowanych do nauki nienadzorowanej prezentujemy tylko zmienne niezależne, które mogą w całości stanowić zbiór uczący, nie podajemy zaś żadnych wytycznych dotyczących rezultatu, jaki chcemy osiągnąć. Klasyfikacje bezwzorcowe służą nam do analizy struktury samych danych i ich podziału na jednorodne zbiory.

Sieci Kohonena są sieciami dwuwarstwowymi. Pierwsza warstwa służy do wprowadzania danych wejściowych;

w drugiej odbywa się przetwarzanie danych. Druga warstwa tworzy tzw. **mapę topologiczną**, a poszczególne neurony stanowią węzły dwuwymiarowej siatki (ryc. 1). Mapa topologiczna przedstawiająca wyuczoną odpowiedź SOM na prezentowane dane wejściowe może mieć różne rozmiary, które są uzależnione od liczby cech charakteryzujących badane zjawiska lub obiekty oraz od liczby analizowanych przypadków.

Sieci Kohonena uczą się struktury danych, a w konsekwencji mogą klasyfikować dane wejściowe w skupienia podobnych klas danych, opierając się na podobieństwie tych danych do siebie (np. cech biometrycznych rostrów belemnitów), reprezentowanych jako skupienia na dwuwymiarowej mapie topologicznej. Dodatkowo sieci Kohonena dają nam możliwość łatwego kontrolowania wielowymiarowej przestrzeni danych wejściowych i prze-



Ryc. 1. A — architektura zastosowanej sieci Kohonena; 29 neuronów w warstwie wejściowej (liczba neuronów odpowiada liczbie poszczególnych cech biometrycznych rostrów); $9 \times 9 = 81$ neuronów w warstwie wyjściowej = mapa topologiczna; B — sposób połączenia neuronów; każdy neuron z warstwy wejściowej jest połączony z 81 neuronami warstwy wyjściowej; C — opisana mapa topologiczna; poszczególne litery charakteryzują centra skupień podobnych danych wejściowych (= cech rostrów, a tym samym podobnych rostrów); obszary zaznaczone kolorem szarym obejmują skupienia rostrów poddane przykładowym analizom

Fig. 1. A — structure of the applied self-organized map (SOM); 29 neurons in the input layer (the number of neurons reflects the number of biometric features used to characterize each guard); $9 \times 9 = 81$ neurons in the output layer = self-organized map; B — the way of joining between the neurons; each input neuron from the input layer is connected with 81 output neurons from the output layer; C — described SOM; individual letters characterize the concentration centers of similar input data (= features of belemnite guard, consequently similar guard); grey areas comprise concentrations of rostra used in further examples of analyses

tworzenia jej w prostą i intuicyjną do analizy przestrzeni dwuwymiarową — mapę topologiczną, która uwzględnia złożone powiązania oraz nieliniowe zależności pomiędzy poszczególnymi danymi uczącymi.

Ważnym aspektem jest fakt, że osoba konstruująca sieć nie daje wskazówek, które z parametrów opisujących dany przypadek są najważniejsze. W trakcie uczenia samoorganizujące się sieci Kohonena zwykle same znajdują najlepsze kryteria podobieństw, często nieprzewidziane i nieznane wcześniej autorowi badań.

Sieci Kohonena umożliwiają lepsze zrozumienie danych pierwotnych, w rozpatrywanym tu przypadku — wszystkich danych biometrycznych rostrów belemnitów, a tym samym ich efektywniejszą analizę, bez względu na to, czy w kolejnych etapach będzie to analiza nadal wykonywana przez człowieka, czy też tradycyjne metody statystyczne. Wyróżnione w ten sposób zbiory opierają się na logice i strukturze wprowadzanych danych, a nie subiektywnych kryteriach.

Dodatkowo, wyuczone sieci Kohonena mogą być używane do tzw. **detekcji nowości**. Wyuczona sieć Kohonena klasyfikuje dane wejściowe w jednorodne zbiory, jednak gdy na wejściu pojawiają się dane odmienne od wszystkich innych (np. nowy gatunek charakteryzowany odmiennymi cechami biometrycznymi), sieć wyróżni je jako nowość, nie próbując „na siłę” zaklasyfikować ich do wcześniej wyróżnionych zbiorów. Przykładem może być sytuacja, gdy wśród dzisiejszych samochodów pojawi się auto sprzed wieku odbiegające wyglądem od współczesnego samochodu. Teoretycznie ma wszystkie jego atrybuty — cztery koła, dwie osie, silnik, kabinę, szyby itd., jednak wyraźnie różni się od jakiegokolwiek grupy („gatunku”) dzisiejszych aut. Na podobnej zasadzie SOM dokonują detekcji nowości w analizowanych danych.

Uczenie sieci Kohonena

Zastosowany algorytm iteracyjny do nauki sieci Kohonena modyfikuje neurony w taki sposób, by neurony i reprezentowane przez nie centra skupień (neurony sąsiedzkie), które są położone blisko siebie w przestrzeni wejść, były rozmieszczane blisko siebie na mapie topologicznej. Wielowymiarowa hiperprzestrzeń danych wejściowych (np. cechy biometryczne rostrów belemnitów) jest przetwarzana automatycznie na dwuwymiarową mapę topologiczną. Umożliwia to w rezultacie intuicyjne interpretowanie wyników.

Działanie algorytmu iteracyjnego w samoorganizujących się mapach Kohonena:

- 1) Inicjalizacja sieci Kohonena w epoce $t = 0$; każdemu neuronowi w warstwie wyjściowej (mapie topologicznej) w sposób losowy przypisywane są wagi — W_{ij} (w zakresie od 0 do 1), tworzące losowe skupienia.
- 2) Prezentowany jest przypadek uczący.
- 3) Obliczana jest odległość d_i (najczęściej wyrażona jako odległość euklidesowa) między każdym z neuronów warstwy wyjściowej a wektorem wag neuronów wejściowych (np. wprowadzanych danych biometrycznych), zgodnie z regułą:

$$d_j(t) = \sum_i^n [x_i(t) - W_{ij}(t)]^2$$

gdzie: $x_i(t)$ jest i -tym komponentem z n -wymiarowej przestrzeni wektorów wejściowych; $W_{ij}(t)$ jest miarą siły

połączeń między i -tym neuronem warstwy wejściowej oraz j -tym neuronem z warstwy wyjściowej w epoce t .

- 4) Wybierany jest tzw. **neuron zwycięski**, dla którego obliczona odległość z pkt. 3 okazała się najmniejsza.
- 5) Następuje aktualizacja wartości wag neuronu zwycięskiego oraz neuronów sąsiednich, zgodnie z formułą:

$$\Delta W_{ij}(t+1) = W_{ij} + \eta(t) \times (x_i(t) - W_{ij}(t)) \times S_j(t)$$

gdzie: $\eta(t)$ — współczynnik uczenia w epoce t ; $S_j(t)$ — funkcja sąsiedztwa w epoce t .

Współczynnik uczenia i promień sąsiedztwa mogą być różne w czasie (w kolejnych epokach) i zmieniają się wtedy od zadanych wartości początkowych do końcowych.

- 6) Kolejna epoka t ; prezentacja kolejnych przypadków uczących i modyfikacja parametrów sieci jak w pkt. 5.
- 7) Jeśli $t < t_{max}$, powrót do pkt. 2, aż do czasu przebiegu całej zadanej liczby epok działania algorytmu uczącego (np. 1000 lub 50 000 epok).

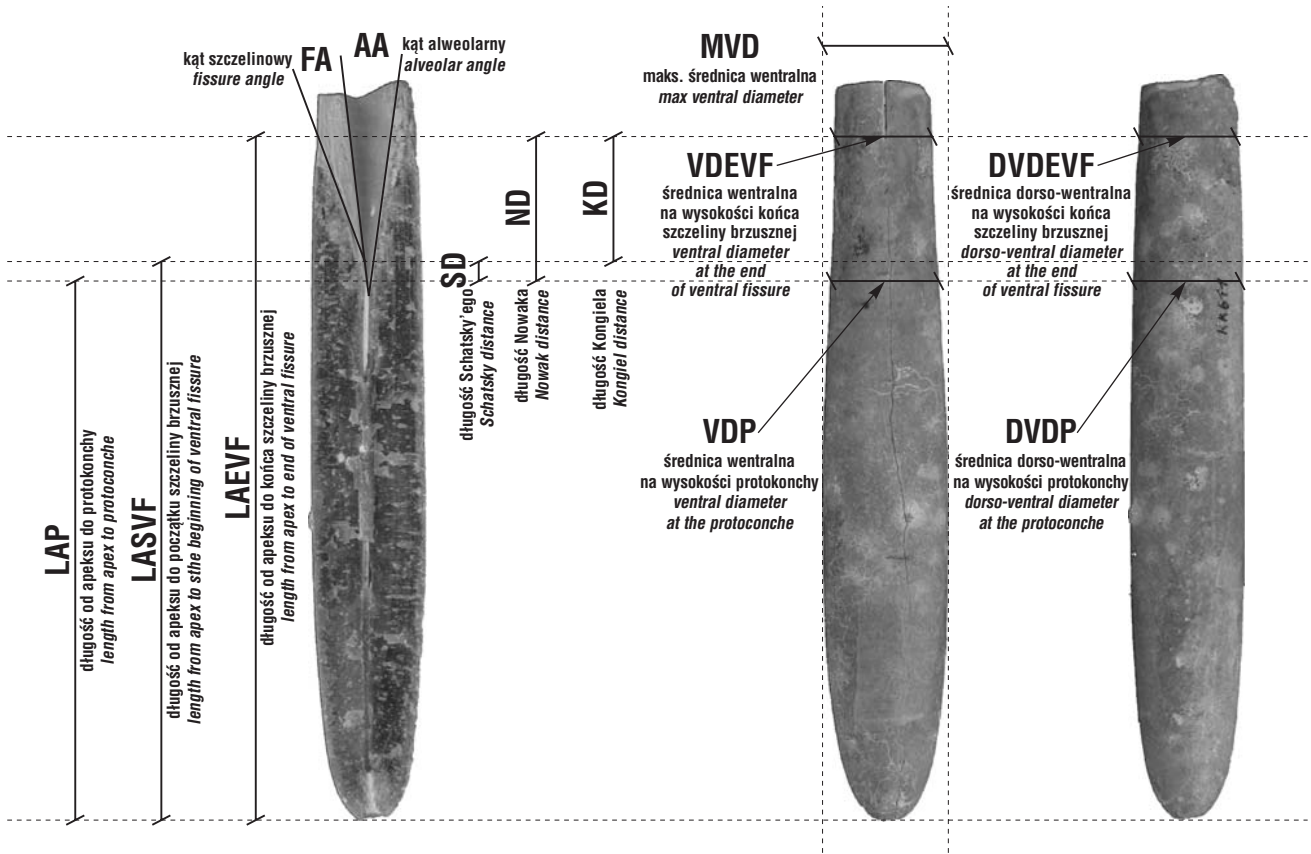
W trakcie każdej epoki działania algorytmu poszczególne przypadki (np. dane biometryczne rostrów belemnitów) prezentowane są sieci Kohonena. Wyznaczany jest neuron zwycięski, czyli neuron najbardziej podobny do wprowadzanych zmiennych wejściowych (danych biometrycznych konkretnego rostrum). W dalszej kolejności model przypadku (rostrum) w neuronie zwycięskim jest modyfikowany, podobnie jak są modyfikowane modele w neuronach sąsiadujących z neuronem zwycięskim, i upodabniane do aktualnego przypadku uczącego.

Dobrze wytrenowana sieć Kohonena tworzy mapę topologiczną odzwierciedlającą wzajemne położenie badanych obiektów na podstawie wprowadzanych danych wejściowych (np. cech biometrycznych rostrów; ryc. 1, zob. też ryc. 2, 3). Podobnym obiektom (= rostrom) konsekwentnie będą odpowiadać te same lub sąsiednie neurony (np. grupa G; ryc. 1C), a podobne grupy obiektów (= morfogrupy A–I) będą reprezentowane na mapie topologicznej jako sąsiednie centra skupień (ryc. 1). Zatem wyróżnione grupy rostrów A–I leżące na mapie topologicznej blisko siebie będą podobne (np. grupy F i G; ryc. 1C; patrz też ryc. 4, 5), a grupy leżące daleko na mapie topologicznej będą różne (np. grupy F i D; ryc. 1C; patrz też ryc. 4, 6).

Zastosowanie sieci Kohonena na przykładzie późnokredowych belemnitów

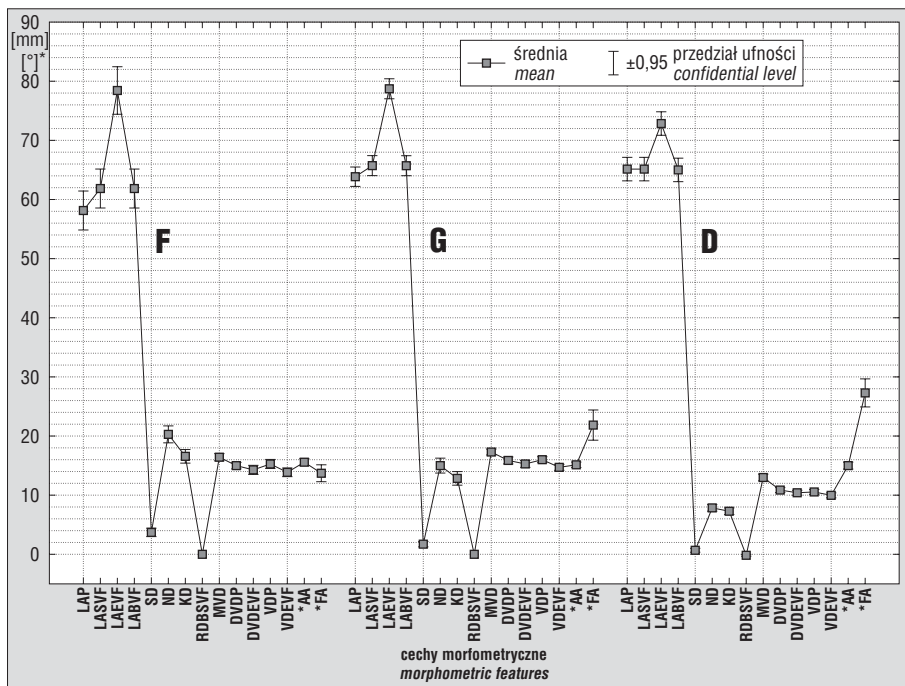
Badane w niniejszej pracy belemnity są zaliczane do podgromady pochwokców (Coleoidea) obejmującej zwierzęta posiadające szkielet wewnętrzny otoczony płaszczem. Są to kopalne głowonogi należące do dziesięciornic i obejmują wyłącznie formy wymarłe. Szkielet wewnętrzny belemnitów — **rostrum** (popularnie, acz błędnie nazywany strzałką piorunową) jest zwykle jedyną pozostałością tych zwierząt, która zachowała się w osadach wieku późnokredowego (ryc. 2). Wydaje się, iż pędziły one tryb życia podobny do przedstawicieli współczesnych głowonogów — kalmarów (Teuthoida) lub mątw (Sepioida).

Podstawą symulacji komputerowych były zebrane przez autora rostra belemnitów z profilu doliny środkowej Wisły obejmującego interwał górnego kampanu oraz dolnego mastrychtu (Remin, 2007). Do testów wytypowano materiał pochodzący z profilu Krons Moor w północno-zachodnich Niemczech zdeponowany w Instytucie



Ryc. 2. Widok ogólny rostrum belemnita oraz podstawowe cechy biometryczne (Remin, 2007); dwie cechy nieilustrowane: LABVF — długość od apeksu do najbardziej cofniętej ku tyłowi części dna szczeliny brzusznej; RDBSVF — długość cofnięcia dna szczeliny brzusznej (= różnica między LASVF a LABVF), szczegóły (Remin, 2007)

Fig. 2. General view of belemnite guard and primary biometric features (Remin, 2007); two not illustrated features: LABVF — length from apex to most back part of ventral fissure; RDBSVF — relative distance of backing of ventral fissure (= LASVF – LABVF), for details see (Remin, 2007)



Ryc. 3. Wykres wartości średnich (wraz z 95% przedziałami ufności dla średnich) 15 cech morfometrycznych charakteryzujących każde rostrum w wyróżnionych grupach morfologicznych F, G, D należących do rodzaju *Belemnella*; objaśnienia skrótów na ryc. 2

Fig. 3. Plot of mean values together with 95% confidential level for 15 direct morphometric features that characterize each belemnite rostra in the distinguished morphogroups F, G, D of the genus *Belemnella*; abbreviations explanation, see Fig. 2

Geologiczno-Paleontologicznym Uniwersytetu w Hamburgu. Materiał ten wybrano z kilku powodów: 1) jest bardzo liczny; 2) okazy są precyzyjnie zlokalizowane w profilu — z dokładnością do kilku centymetrów; 3) eksponaty stanowią oryginalną kolekcję belemnitów Schulza, na podstawie której został stworzony obowiązujący do dziś schemat biostratygraficzny dolnego mastrychtu (Schulz, 1979); 4) wyniki badań kolekcji testowej klasycznymi metodami rozpoznania taksonomicznego można porównać z rezultatami osiągniętymi dzięki wykorzystaniu SOM. Ponadto w celach porównawczych wykorzystano kolekcję Kongiela przechowywaną w Muzeum Ziemi PAN w Warszawie, a także bogate kolekcje brytyjskie znajdujące się w Muzeum Historii Naturalnej w Londynie oraz w siedzibie Brytyjskiej Służby Geologicznej w Keyworth.

Tab. 1. Statystyki opisowe analizowanej grupy F
Table 2. Descriptive statistics of the F group

Zmienne Variables	Liczba obserwacji Number of observations	Średnia Mean	Ufność Confidential interval -95%	Ufność Confidential interval +95%	Mediana Median	Minimum Minimum	Maksimum Maximum	Odchylenie stand. Standard deviation	Błąd stand. Standard error	Wariancja Variance
[mm]										
LAP	14	58,11	54,80	61,43	58,43	45,82	65,86	5,740	1,534	32,94
LASVF	14	61,80	58,50	65,10	61,77	49,78	69,59	5,716	1,528	32,67
LAEVF	14	78,41	74,40	82,42	79,05	66,71	87,42	6,947	1,857	48,26
LABVF	14	61,80	58,50	65,10	61,77	49,78	69,59	5,716	1,528	32,67
SD	14	3,69	3,01	4,36	3,45	2,29	7,04	1,173	0,313	1,37
ND	14	20,30	18,90	21,70	20,58	15,41	23,92	2,432	0,650	5,91
KD	14	16,61	15,44	17,78	17,03	12,63	19,23	2,029	0,542	4,12
RDBSVF	14	–	–	–	–	–	–	–	–	–
MVD	14	16,41	15,83	16,99	16,79	14,56	17,88	1,002	0,268	1,00
DVDP	14	14,95	14,42	15,47	14,83	13,17	16,29	0,917	0,245	0,84
DVDEVF	14	14,31	13,62	15,00	13,80	12,64	16,44	1,193	0,319	1,42
VDP	14	15,34	14,69	15,99	15,25	13,69	17,23	1,123	0,300	1,26
VDEVF	14	13,85	13,20	14,51	13,54	12,26	15,77	1,132	0,303	1,28
[°]										
AA	14	15,59	15,04	16,14	15,50	14,03	17,52	0,956	0,255	0,91
FA	14	13,71	12,22	15,19	13,29	9,92	18,56	2,570	0,687	6,60

Tab. 2. Statystyki opisowe analizowanej grupy G
Table 2. Descriptive statistics of the G group

Zmienne Variables	Liczba obserwacji Number of observations	Średnia Mean	Ufność Confidential interval -95%	Ufność Confidential interval +95%	Mediana Median	Minimum Minimum	Maksimum Maximum	Odchylenie stand. Standard deviation	Błąd stand. Standard error	Wariancja Variance
[mm]										
LAP	22	63,81	62,19	65,43	64,10	55,80	69,95	3,653	0,779	13,35
LASVF	19	65,67	63,93	67,41	66,24	59,02	71,61	3,608	0,828	13,02
LAEVF	21	78,71	77,02	80,39	79,00	71,88	85,38	3,703	0,808	13,71
LABVF	19	65,67	63,93	67,41	66,24	59,02	71,61	3,608	0,828	13,02
SD	20	1,73	1,28	2,19	1,87	0,00	3,21	0,972	0,217	0,94
ND	22	14,99	13,78	16,20	15,12	10,12	21,11	2,739	0,584	7,50
KD	19	12,88	11,79	13,98	13,08	9,01	16,51	2,267	0,520	5,14
RDBSVF	23	–	–	–	–	–	–	–	–	–
MVD	22	17,33	16,92	17,74	17,30	15,37	19,23	0,932	0,199	0,87
DVDP	23	15,91	15,41	16,41	15,96	13,10	17,67	1,162	0,242	1,35
DVDEVF	22	15,23	14,60	15,85	15,37	11,75	17,80	1,413	0,301	2,00
VDP	23	16,05	15,61	16,49	16,06	13,19	18,09	1,011	0,211	1,02
VDEVF	22	14,70	14,19	15,20	14,67	11,36	16,51	1,143	0,244	1,31
[°]										
AA	20	15,15	14,69	15,62	15,20	12,57	16,69	0,996	0,223	0,99
FA	19	21,85	19,21	24,49	21,54	12,92	34,65	5,476	1,256	29,98

Do prowadzonych badań i symulacji komputerowych została stworzona baza danych, obejmująca pomiary 15 cech morfometrycznych (ryc. 2) i 14 proporcji opartych na tych cechach, około 750 okazów (macierz danych obejmująca około 21 000 rekordów pomiarowych). W anali-

zach wykorzystano program *Statistica Neural Networks PL*. Z wielu wykonanych symulacji uczenia SOM w niniejszym artykule zaprezentowano jedną, na podstawie której scharakteryzowano podstawowe możliwości SOM w klasyfikacji organizmów kopalnych.

→

Ryc. 4. *Belemnella* sp. F; grupa F na mapie topologicznej (ryc. 1); porównaj również ryc. 3 oraz statystyki opisowe (tab. 1); pary — np. 1a (widok ogólny) oraz 1b (rozłupane rostrum z widoczną strukturą wewnętrzną) reprezentują jeden okaz

Fig. 4. *Belemnella* sp. F; group F on the SOM (Fig. 1); compare also Fig. 3 and descriptive statistics (Table 1); pairs i.e. 1a (general view) and 1b (crushed guard showing internal structure) represent the same guard

→

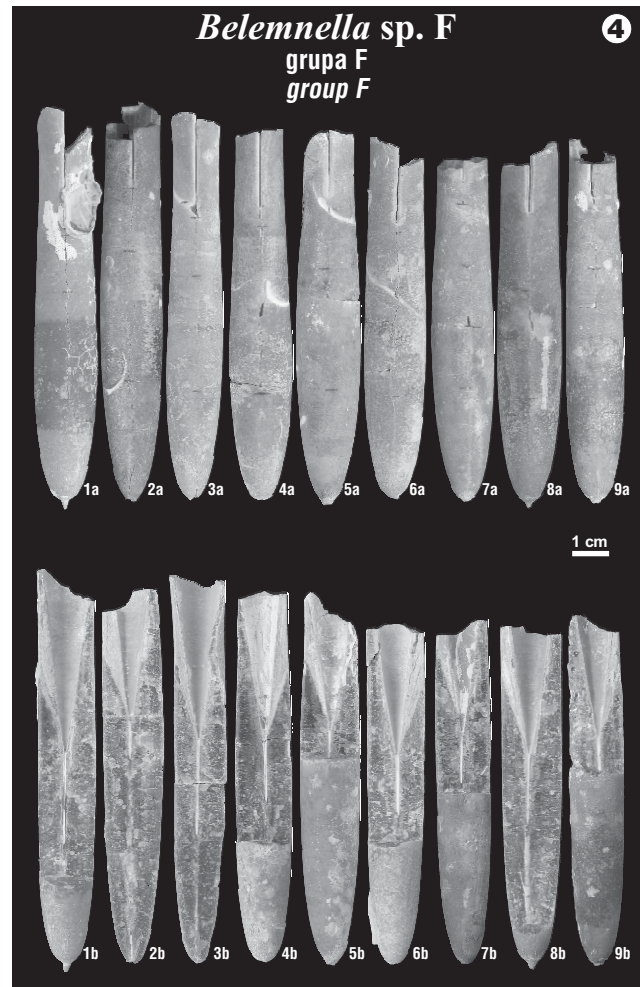
Ryc. 5. *Belemnella* sp. G; grupa G na mapie topologicznej (ryc. 1); porównaj również ryc. 3 oraz statystyki opisowe (tab. 2); pary — np. 1a (widok ogólny) oraz 1b (rozłupane rostrum z widoczną strukturą wewnętrzną) reprezentują jeden okaz

Fig. 5. *Belemnella* sp. G; group G on the SOM (Fig. 1); compare also Fig. 3 and descriptive statistics (Table 2); pairs i.e. 1a (general view) and 1b (crushed guard showing internal structure) represent the same guard

→

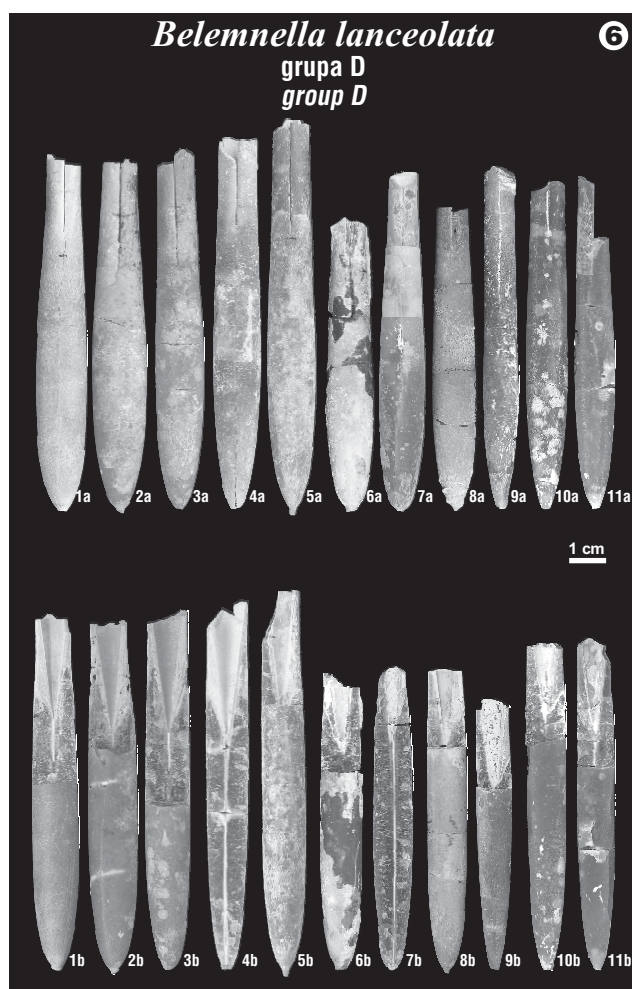
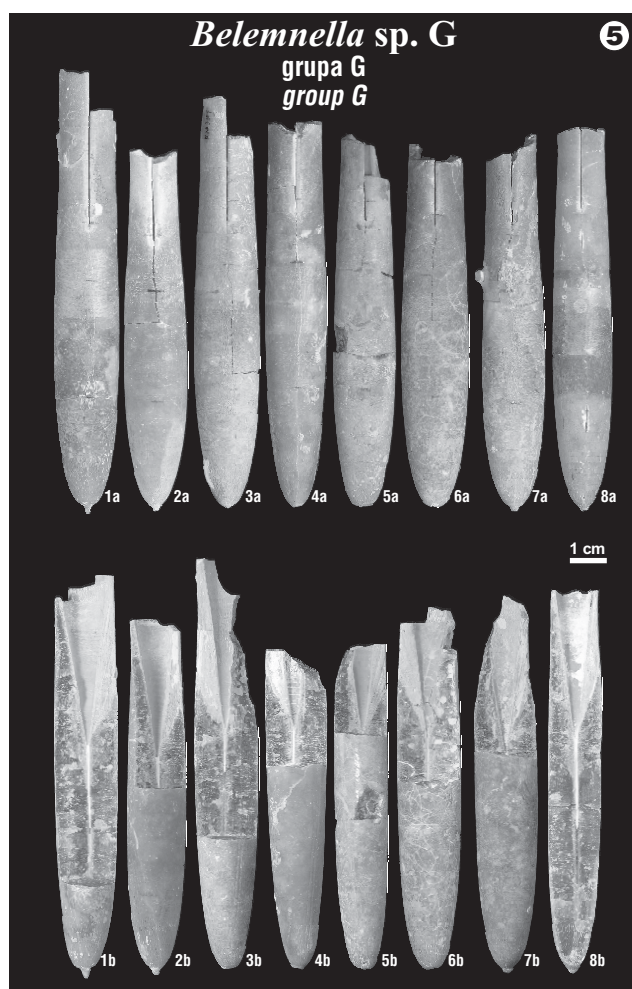
Ryc. 6. *Belemnella lanceolata*; grupa D na mapie topologicznej (ryc. 1); porównaj również ryc. 3 oraz statystyki opisowe (tab. 3); pary — np. 1a (widok ogólny) oraz 1b (rozłupane rostrum z widoczną strukturą wewnętrzną) reprezentują jeden okaz

Fig. 6. *Belemnella lanceolata*; group D on the SOM (Fig. 1); compare also Fig. 3 and descriptive statistics (Table 3); pairs i.e. 1a (general view) and 1b (crushed guard showing internal structure) represent the same guard



Tab. 3. Statystyki opisowe analizowanej grupy D
Table 3. Descriptive statistics of the D group

Zmienne Variables	Liczba obserwacji Number of observations	Średnia	Ufność	Ufność	Mediana	Minimum	Maksimum	Odchylenie stand. Standard deviation	Błąd stand. Standard error	Wariancja
		Mean	Confidential interval -95%	Confidential interval +95%						
[mm]										
LAP	61	65,15	63,10	67,19	64,78	47,68	86,16	7,985	1,022	63,76
LASVF	47	65,16	63,19	67,12	65,36	48,86	78,22	6,692	0,976	44,79
LAEVF	60	72,89	70,88	74,89	73,74	54,07	90,37	7,770	1,003	60,37
LABVF	47	64,96	63,05	66,86	65,31	48,86	77,29	6,496	0,948	42,20
SD	53	0,73	0,44	1,01	0,00	0,00	4,61	1,023	0,141	1,05
ND	65	7,88	7,36	8,40	7,87	3,52	14,14	2,090	0,259	4,37
KD	52	7,35	6,78	7,92	6,98	3,52	12,68	2,036	0,282	4,14
RDBSVF	68	-0,14	-0,23	-0,05	0,00	-2,13	0,00	0,375	0,046	0,14
MVD	66	12,98	12,43	13,53	13,17	7,18	17,17	2,224	0,274	4,95
DVDP	67	10,88	10,52	11,25	11,14	6,81	13,23	1,492	0,182	2,23
DVDEVF	66	10,37	10,03	10,70	10,56	6,77	12,84	1,368	0,168	1,87
VDP	67	10,56	10,17	10,95	10,74	5,98	13,18	1,580	0,193	2,50
VDEVF	66	9,93	9,58	10,29	10,01	5,83	12,52	1,450	0,179	2,10
[°]										
AA	53	15,04	14,55	15,52	14,98	11,66	19,10	1,764	0,242	3,11
FA	52	27,22	24,90	29,54	25,00	11,39	48,97	8,323	1,154	69,27



W prezentowanym przykładzie (ryc. 1) do analizy zastosowano sieć Kohonena charakteryzującą się 29-wymiarową hiperprzestrzenią danych wejściowych (cechy biometryczne rostrów; ryc. 2). Rozmiar mapy topologicznej został określony na 9×9 neuronów = 81 neuronów w warstwie wyjściowej (ryc. 1). Do symulacji wykorzystano 227 rostrów ze zbioru testowego pochodzącego z profilu Krons Moor w północno-zachodnich Niemczech, które mają pełną charakterystykę biometryczną.

Analiza mapy topologicznej (ryc. 1C) stanowiącej odpowiedź SOM na wprowadzane zmienne wejściowe pozwoliła w badanym przypadku na wyróżnienie 9 skupień, do których sieć zaklasyfikowała testowane rostra. Skupienia na rycinie 1C oznaczono literowo A–I. Oznaczenia „???” (ryc. 1C) reprezentują neurony, których sposób aktywacji i, w konsekwencji, przynależność niektórych rostrów nie były wystarczająco jednoznaczne, by zaklasyfikować je do dobrze zdefiniowanych grup A–I.

Zgodnie z założeniami działania algorytmu uczącego poszczególne skupienia neuronów A–I na mapie topologicznej (ryc. 1C) odzwierciedlają wzajemne powinowactwo morfoprzestrzenne wyróżnionych morfogrów. Czyli morfogrupy leżące na mapie topologicznej blisko siebie będą podobne; grupy leżące na przeciwległych krańcach mapy topologicznej będą różne. Możemy się zatem spodziewać, że wyróżniona grupa F będzie mniej lub bardziej podobna do grupy G, gdyż obydwie te grupy są wobec siebie sąsiedzkie (porównaj ryc. 1C i ryc. 3–5, tab. 1, 2).

Podobnie możemy się spodziewać dużych różnic morfologicznych rostrów belemnitów należących odpowiednio do grup F i D oraz G i D. Wynika to z faktu, iż grupy te leżą na przeciwległych krańcach mapy topologicznej (patrz ryc. 1C i ryc. 3–6, tab. 1–3).

Przedstawione w tym rozdziale założenia dotyczące podobieństwa i różnic między analizowanymi grupami F, G, D zweryfikowano, analizując standardowe wykresy wartości średnich wraz z 95% przedziałami ufności (ryc. 3) poszczególnych cech w wyróżnionych morfogrupach oraz analizując podstawowe statystyki opisowe (tab. 1–3).

Cechy biometryczne rostrów belemnitów należących do morfogrów F oraz G (sąsiedzkie na mapie topologicznej; ryc. 1C oraz ryc. 3, tab. 1, 2) wykazują duże podobieństwo. Zbieżne są wartości cech długościowych oraz ich wzajemna relacja (LAP, LASVF, LAEVF, LABVF), podobne także są wartości średnic (MVD, DVDP, DVDEVF, VDP, VDEVF). Istotne różnice obserwujemy w cechach wewnętrznych (SD, ND, KD) oraz w kącie szczelinowym (FA).

Charakterystyka biometryczna morfogrów F i D oraz G i D (leżące daleko na mapie topologicznej; ryc. 1C oraz ryc. 3, tab. 1–3) świadczy o znaczących różnicach wielkości cech biometrycznych, pomimo relatywnie podobnych wartości cech długościowych w morfogrupach F, G i D (ryc. 3; tab. 1–3). Wzajemna relacja tych cech w grupie D jest całkowicie odmienna niż w grupach F i G (ryc. 3; tab. 1–3). Grupa D różni się ponadto od grup F i G wszyst-

kimi pozostałymi średnimi wartościami cech i są to istotne różnice.

Wykonane analizy, w których wykorzystano standardowe metody statystyczne (ryc. 3 oraz tab. 1–3), pozwoliły pozytywnie zweryfikować rezultaty uzyskane z zastosowaniem SOM. Wyróżnione grupy A–I charakteryzują się morfologiczną spójnością wewnątrzgrupową oraz istotnymi różnicami międzygrupowymi wielu cech badanych rostrów. Różnice między grupami podkreślone są ponadto przez odmienne interwały stratygraficzne, w których są obecne poszczególne morfotypy. Wydaje się zatem uzasadnione traktowanie tych morfogrów jako wydzieleni monospecyficznych, reprezentujących najprawdopodobniej oddzielne gatunki belemnitów.

Analiza materiału testowego — belemnitów, charakteryzujących się bardzo małą dyspersją morfologiczną — z użyciem implementowanej metodyki sztucznych sieci neuronowych Kohonena do bezwzorcowych klasyfikacji potwierdziła wysoką skuteczność i użyteczność tej metody oraz znaczną konkurencyjność w stosunku do metod stosowanych dotychczas (np. Kongiel, 1962; Schulz, 1979; Christensen, 1995). W wyniku wszystkich przeprowadzonych symulacji z zastosowaniem sieci Kohonena w badanym interwale górnego kampanu i dolnego mastrychtu profilu doliny środkowej Wisły wyróżniono 15 grup morfologicznych reprezentujących rodzaje *Belemnitella* oraz *Belemnella* (patrz Remin, 2007).

Podsumowanie i perspektywy

Przeprowadzone badania świadczą o tym, że ANN mają wielki potencjał. Można je stosować do rozwiązywania problemów klasyfikacji taksonomicznej, szczególnie tam, gdzie analizie poddajemy duże ilości danych statystycznych lub biometrycznych. Otwiera to nowe perspektywy w podstawowych badaniach paleontologicznych wielu grup skamieniałości (np. belemnitów, amonitów, inoceramów, ramienionogów, trylobitów, koralowców). ANN dają ponadto możliwość wychwytywania nawet niewielkich zmian i trendów filetycznych i filogenetycznych w obrębie analizowanych grup faunistycznych, co w wielu wypadkach może się przełożyć na wzrost ich potencjału biostratygraficznego.

Perspektywy dalszego wykorzystania ANN wydają się być dobre, zważywszy na fakt, że badacze mogą się posługiwać wieloma darmowymi programami dostępnymi na różnych stronach internetowych. Programy te są projektowane do pracy z większością popularnych systemów operacyjnych. Dodatkowo do dyspozycji badaczy są oddawane profesjonalne pakiety statystyczne obejmujące procedury uczenia ANN. Pozwala to w efekcie na analizę bardzo złożonych problemów badawczych i wizualizację otrzymanych rezultatów z jakością nieosiągalną klasycznymi metodami.

Składam serdeczne podziękowania promotorowi prof. Ireneuszowi Walaszczykowi oraz recenzentom rozprawy doktorskiej — prof. Ryszardowi Marcinowskiemu i prof. Marianowi Gasińskiemu za owocne dyskusje i cenne uwagi, które w rezultacie przyczyniły się do lepszej i czytelniejszej formy niniejszego artykułu. Dziękuję również dr. Andrzejowi Krukowi za weryfikację pierwotnej wersji tekstu, a także recenzentom tego artykułu — prof. Ryszardowi Tadeusiewiczowi oraz dr. hab. Marcinowi Machalskiemu za krytyczne i cenne uwagi.

Badania finansowane były z grantu Ministerstwa Nauki i Informatyzacji nr 2 P04D 020 27, ze środków wydziału Geologii UW oraz ze środków własnych autora.

Literatura

- BALDWIN J.L., OTTE D.N. & WHEATLEY C.L. 1989 — Computer emulation of human mental processes. Application of neural network simulators to problems in well log interpretation. Society of Petroleum Engineers of AIME, (Paper) SPE, OMEGA: 481–493.
- CHANG H.C., KOPASKA-MERKEL D.C. & CHEN H.C. 2002 — Identification of lithofacies using Kohonen self-organizing maps. Computers and Geosciences, 28: 223–229.
- CHON T.S., PARK Y.S., MOON K.H. & CHA E. 1996 — Patternizing communities by using an artificial neural network. Ecological Modelling, 90: 69–78.
- CHON T.S., PARK Y.S. & PARK J.H. 2000 — Determining temporal pattern of community dynamics by using unsupervised learning algorithms. Ecological Modelling, 132: 151–166.
- CHRISTENSEN W.K. 1995 — *Belemnitella* from the Upper Campanian and Lower Maastrichtian Chalk of Norfolk, England. Special Papers in Palaeontology, 51, 1–84.
- FAUSETT L. 1994 — Fundamentals of Neural Networks. Prentice Hall, New York.
- GUÉGAN J.F., LEK S. & OBERDORFF T. 1998 — Energy availability and habitat heterogeneity predict global riverine fish diversity. Nature, 391: 382–384.
- GISKE J., HUSE G. & FIKSEN O. 1998 — Modelling spatial dynamics of fish. Rev. Fish. Biol. Fish. 8: 57–91.
- HAYKIN S. 1994 — Neural Networks: A Comprehensive Foundation. MacMillan Publishing, New York.
- KAMINSKAS D. & MALMGREN B.A. 2004 — Comparison of pattern-recognition techniques for classification of Silurian rocks from Lithuania based on geochemical data. Norsk Geologisk Tidsskrift, 84: 117–124.
- KOHONEN T. 1982 — Self-organized formation of topologically correct feature maps. Biological Cybernetics, 43: 59–69.
- KONGIEL R. 1962 — On belemnites from the Maastrichtian, Campanian and Santonian sediments in the Middle Vistula Valley (Central Poland). Pr. Muz. Ziemi, 5: 1–148.
- KRUK A. 2007 — Role of habitat degradation in determining fish distribution and abundance along the lowland Warta River, Poland. J. Appl. Ichthyol., 23: 9–18.
- MALMGREN B.A. & NORDLUND U. 1996 — Application of artificial neural networks to chemostratigraphy. Paleoceanogr., 11: 505–512.
- MALMGREN B.A. & NORDLUND U. 1997 — Application of artificial neural networks to paleoceanographic data. Palaeogeogr., Palaeoclimatol., Palaeoecol., 136: 359–373.
- MARMO R., AMODIO S. & CANTONI V. 2006 — Microfossils shape classification using a set of width values. Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06). Vol. 1, 20–24 Aug., 2006: 691–694.
- PATTERSON D. 1996 — Artificial Neural Networks. Prentice Hall, Singapore.
- RECKNAGEL F., FRENCH M., HARKONEN P. & YABUNAKA K.I. 1997 — Artificial neural network approach for modelling and prediction of algal blooms. Ecological Modelling, 96: 11–28.
- REMIN Z. 2007 — Analiza paleontologiczna i znaczenie stratygraficzne belemnitów górnego kampanu i dolnego mastrychtu profilu doliny środkowej Wisły. Arch. WG UW, Warszawa.
- ROGERS S.J., FANG J.H., KARR C.L. & STANLEY D.A. 1992 — Determination of lithology from well logs using a neural network. Am. Assoc. Pet. Geol. Bull., 76: 731–739.
- SCHULZ M.G. 1979 — Morphometrisch-variationsstatistische Untersuchungen zur Phylogenie der Belemniten-Gattung *Belemnella* im Untermaastricht NW-Europas. Geol. Jahrb., A47: 3–157.
- SCARDI M. 1996 — Artificial neural networks as empirical models for estimating phytoplankton production. Marine Ecol. Progr. Series 139: 289–299.
- SEGNER I., BOULARD T. & BAILEY B.J. 1994 — Neural network models of the greenhouse climate. J. Agric. Eng. Res. 59: 203–216.
- StatSoft 2001 — Statistica Neural Networks PL. (I) Wprowadzenie do sieci neuronowych; (II) Poradnik użytkownika; (III) Przewodnik problemowy. StatSoft Polska Sp. z o.o., Kraków.
- StatSoft 2006 — Elektroniczny Podręcznik Statystyki PL. Kraków. WEB: <http://www.statsoft.pl/textbook/stathome.html>.

Praca wpłynęła do redakcji 19.10.2007 r.

Po recenzji akceptowano do druku 17.12.2007 r.