

1

DOBÓR ARCHITEKTURY SIECI NEURONOWEJ WYKORZYSTYWANEJ DO OPISU ZMIAN WIELKOŚCI SEJSMICZNOŚCI INDUKOWANEJ

1.1 WPROWADZENIE

Większość rejestrowanych wstrząsów w kopalniach GZW jest związanych bezpośrednio z prowadzonymi robotami górniczymi. Ogniska tych wstrząsów przemieszczają się wraz z postępowaniem frontów ścianowych oraz przodków drążonych wyrobisk korytarzowych; aktywność sejsmiczna jest silnie zależna od lokalnych warunków geologicznych, a także parametrów prowadzonych robót górniczych. Parametry te (m.in. głębokość zalegania złoża, wysokość furty eksploatacyjnej, sposób kierowania stropem, odległość od resztek i krawędzi wytworzonych w sąsiednich pokładach) także istotnie wpływają na obliczane zmiany wartości energii właściwej w górotworze. Do szacowania zmian energetycznych zachodzących w wytrzymałych warstwach skalnych autorzy wykorzystują metodę analityczną, opierającą się na rozwiązaniu przemieszczeniowego zadania brzegowego liniowej teorii sprężystości podanym przez H. Gila [2], określającym rozkład naprężeń i odkształceń w półprzestrzeni wokół pustki o kształcie prostokąta. Obliczone wartości składowych tensorów naprężenia i odkształcenia pozwalają określić energię właściwą odkształcenia sprężystego, zgodnie ze wzorem Clapeyrona:

$$\Phi = \frac{1}{2} T_{\sigma} T_{\varepsilon}, \quad (1.1)$$

gdzie:

Φ – energia właściwa odkształcenia sprężystego, J/m³,

T_{σ} – tensor stanu naprężenia, MPa,

T_{ε} – tensor stanu odkształcenia.

Podczas prowadzenia robót górniczych, obliczane wartości energii właściwej w podbieranych (a także nadbieranych) warstwach skalnych ulegają dużym zmianom, których przebieg można wiązać ze zmianami poziomu sejsmiczności indukowanej. Ponieważ obserwuje się zróżnicowanie liczby i wydatku energetycznego wstrząsów występujących na wybiegach oraz w zrobach prowadzonych wyrobisk ścianowych, w prowadzonych badaniach uwzględniane są osobno, obliczane w wytypowanych,

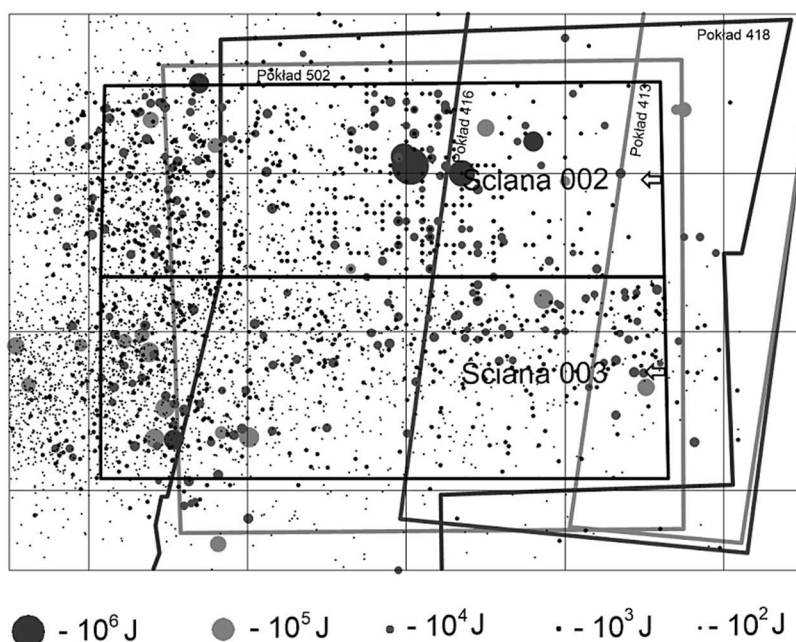
wstrząsogennych warstwach skalnych, przyrosty oraz spadki energii właściwej. Dodatkowo, w celu uwzględnienia ewentualnego zniszczenia struktury skał, które zaszło wskutek przeprowadzenia wcześniejszych robót górniczych, powyższe przyrosty oraz spadki energii właściwej rozdzielane są na zachodzące przed hipotetycznym zniszczeniem skał (wskaźnik wyężenia $W < 1$) oraz po nim ($W \geq 1$). Oznacza to, że zmiany energetyczne zachodzące w danej warstwie wstrząsogennej opisywane są czterema wskaźnikami.

W pracy [1] wskazano na możliwość wykorzystania do prognozy zmian poziomu sejsmiczności indukowanej odpowiednio wytrenowanej sieci neuronowej. Do nauki sieci wykorzystano dane o zmianach energetycznych zachodzących w wytrzymałych, wstrząsogennych warstwach skalnych oraz o poziomie sejsmiczności indukowanej opisywanym poprzez określanie gęstości energii wstrząsów [J/m^3].

W artykule przedstawiono wyniki dalszych badań mających na celu ustalenie optymalnej architektury sieci neuronowej wykorzystywanej do prognozy zmian wielkości gęstości energii wstrząsów indukowanych prowadzonymi robotami górniczymi. W trakcie analiz wykorzystano dane pochodzące z silnie zagrożonego wstrząsami i tąpniętami rejonu robót górniczych prowadzonych w jednej z kopalń węgla kamiennego GZW.

1.2 CHARAKTERYSTYKA REJONU BADAŃ

Przedmiotowa eksploatacja była prowadzona w polach dwóch ścian: 002 i 003 w górnej warstwie pokładu 504 (rys. 1.1), którego miąższość zmienia się od 2,8 do 7,0 m (wysokość furty eksploatacyjnej wynosiła do 3,4 m). Głębokość zalegania wynosi około 850 m, upad warstw $4^\circ \div 16^\circ$ w kierunku południowym.



Rys. 1.1 Kontury ścian 002 i 003 w pokładzie 504 z naniesionymi krawędziami eksploatacji dokonanej w pokładach 416, 418 i 502 oraz epicentrami wstrząsów

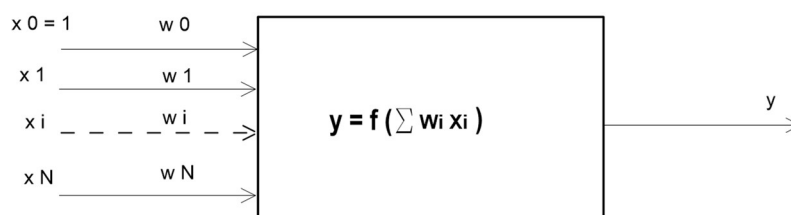
W rejonie omawianych wyrobisk ścianowych wyeksploatowano wyżej zalegające pokłady, m.in.: 413 w odległości około 270 m, 416 w odległości około 160 m, 418 w odległości około 130 m i 502 w odległości około 70 m od pokładu 504. W profilu geologicznym występuje kilka potencjalnie wstrząsogennych warstw piaskowca. W wybranych czterech z nich obliczono zmiany energii właściwej, zachodzące w trakcie prowadzenia ścian 002 i 003. Są to warstwy o miąższościach około: 30 m, 31 m, 23 m i 20 m, zalegające w odległościach odpowiednio: 450 m, 350 m, 100 m i 20 m nad pokładem 504.

Rozpatrywana eksploatacja pokładu 504 indukowała 26 wysokoenergetycznych wstrząsów górotworu: 20 o energii rzędu 10^5 J oraz 6 o energii rzędu 10^6 J. Ponadto, zarejestrowano 194 zjawiska o energii rzędu 10^4 J. Ogółem wystąpiło 5000 wstrząsów o sumarycznej energii około 4×10^7 J. Najsilniejszy wstrząs wystąpił w trakcie prowadzenia ściany 002 i osiągnął energię 6×10^6 J. Podobnie jak pozostałe, nie spowodował on wystąpienia skutków w wyrobiskach górniczych.

1.3 PODSTAWOWE INFORMACJE O SIECIACH NEURONOWYCH

Sieci neuronowe stanowią uniwersalny układ aproksymacyjny, pozwalający odwzorowywać wielowymiarowe zbiory danych, mający zdolność uczenia się i adaptacji do zmieniających się warunków środowiskowych, zdolność uogólniania nabytej wiedzy, stanowiąc pod tym względem system sztucznej inteligencji [3]. Powstały one w wyniku badań, które dotyczyły budowy modeli podstawowych struktur występujących w mózgu. Mózg człowieka składa się z bardzo dużej liczby (około 10 miliardów) elementarnych komórek nerwowych – neuronów połączonych ze sobą w formie skomplikowanej sieci. Na jeden neuron przypada przeciętnie kilka tysięcy połączeń. Neuron w uproszczeniu posiada wiele wejść informacyjnych, całość, scalające sygnały z tych wejść oraz pojedyncze włókno przewodzące informację wyjściową. Pobudzony przez sygnały wejściowe neuron przechodzi do stanu aktywnego, gdy łączny docierający sygnał przekroczy poziom progowy. W stanie aktywnym wysyła on sygnał do połączonych z nim kolejnych neuronów poprzez złącza, modyfikujące sygnały i stanowiące nośnik pamięci. Proces uczenia się polega m.in. na wprowadzaniu zmian w parametrach połączeń pojedynczych neuronów.

Sztuczne sieci neuronowe korzystają z bardzo prostego modelu neuronu (rys. 1.2), w którym przyjęto, że obliczana jest ważona suma sygnałów wejściowych i gdy przekroczy ona poziom progowy, sygnał podawany jest na wyjście.

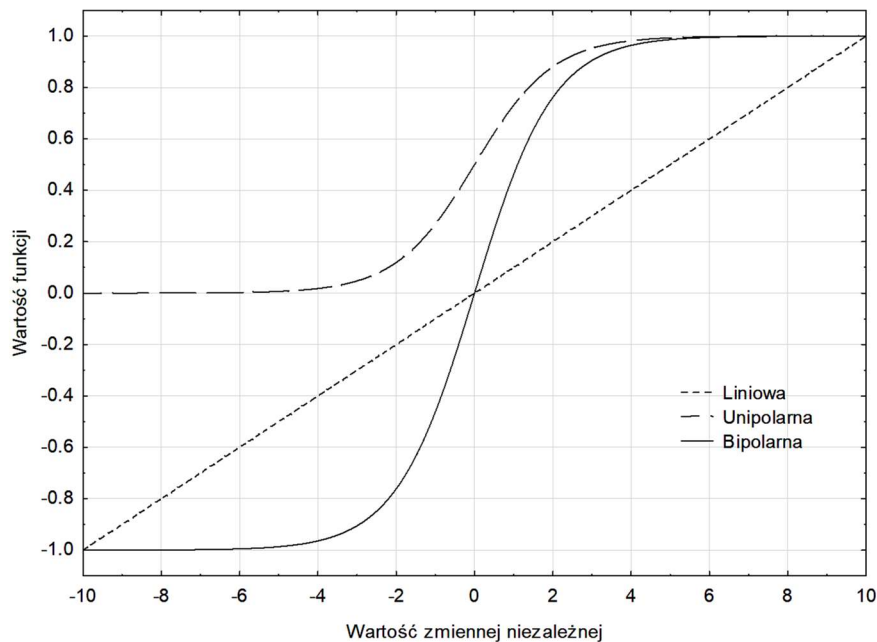


Rys. 1.2 Model neuronu

Wielkość sygnału wyjściowego obliczana jest za pomocą pewnej funkcji f .

Funkcja aktywacji f może mieć różną postać, w prowadzonych badaniach uwzględniano następujące: liniową, logistyczną (unipolarną – w największym stopniu przypominającą zachowanie rzeczywistego neuronu) oraz tangens hiperboliczny (bipolarną – korzystniejszą od unipolarnej w trakcie nauki sieci neuronowej) – rys. 1.3.

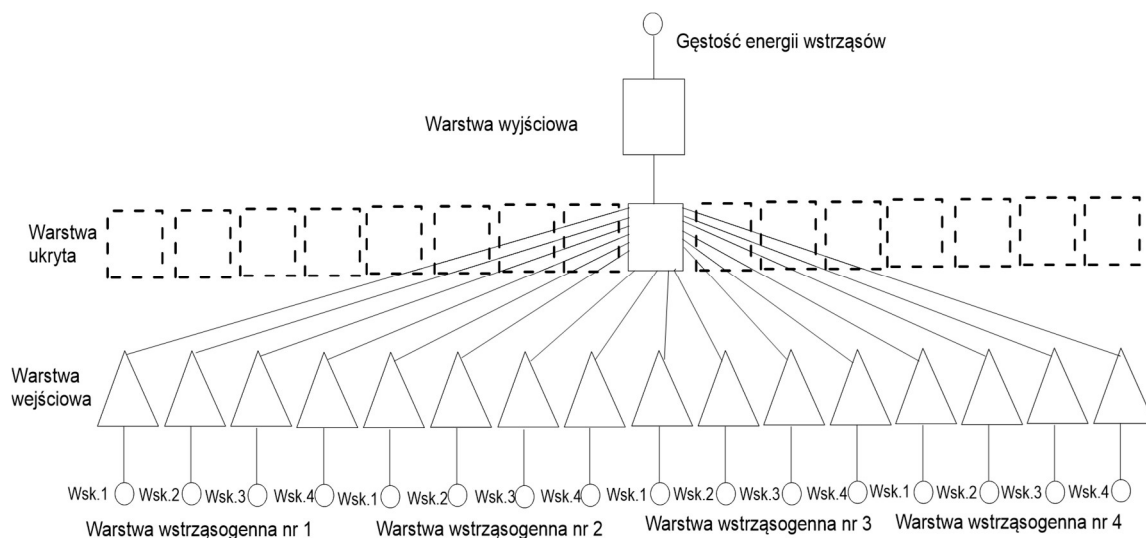
W odróżnieniu od metod parametrycznych (np. metody analizy regresji), wymagających znajomości funkcji wiążących zmienne niezależne ze zmiennymi zależnymi, sieci neuronowe, będące metodami nieparametrycznymi, nie zakładają posiadania takiej apriorycznej informacji – automatycznie uczą się na podanych przez użytkownika danych pokazujących, jak manifestuje się interesująca zależność. Sieci neuronowe mają wyjątkową zdolność znajdowania sensu i znaczenia, reguł i trendów w skomplikowanych strukturach zaszumionych i nieprecyzyjnych danych [5].



Rys. 1.3 Zastosowane charakterystyki neuronów

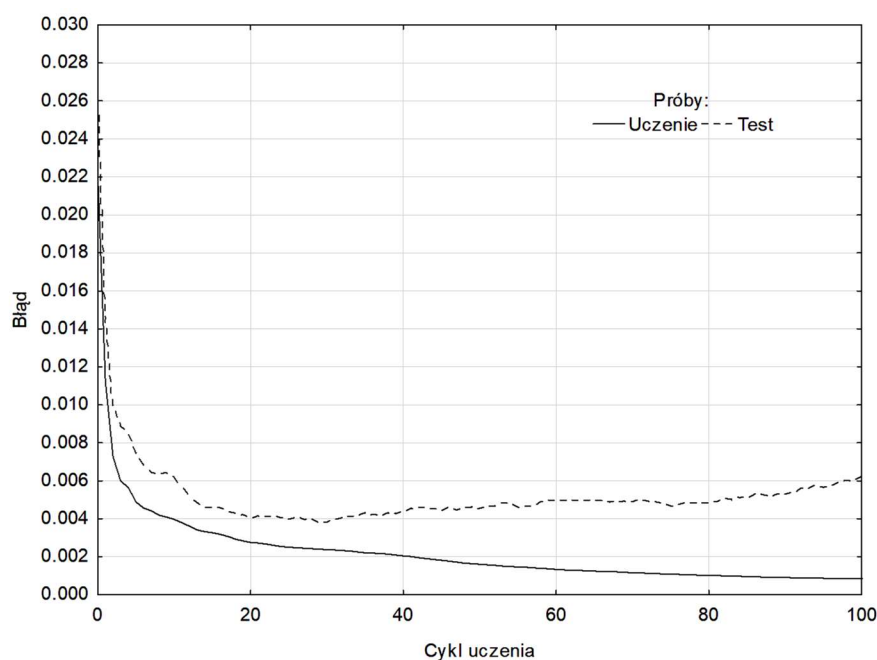
Możliwość różnego sposobu połączenia neuronów między sobą i ich współdziałania spowodowała powstanie różnych typów sieci neuronowych i związanych z nimi metod doboru wag (uczenia sieci). W zagadnieniach regresyjnych najczęściej wykorzystywane są sieci jednokierunkowe, wielowarstwowe (rys. 1.4).

Wartości zmiennych wejściowych przekazywane są do neuronów wejściowych, następnie do neuronów warstw ukrytych i warstwy wyjściowej. W każdym neuronie obliczana jest ważona suma wejść. Następnie wyznaczana jest wartość wyjściowa, zależna od informacji wejściowej i przyjętej funkcji aktywacji. Wynik działania całej sieci pojawia się na wyjściach neuronów warstwy wyjściowej.



Rys. 1.4 Model sieci jednokierunkowej, wielowarstwowej

Do nauki sieci (doboru wag dla każdego z wejść neuronów) najczęściej wykorzystywany jest algorytm wstecznej propagacji błędów. Przed rozpoczęciem nauki zbiór danych jest dzielony na dwa podzbiory: liczniejszy, danych uczących oraz danych testowych. Pierwszy z powyższych zbiorów jest wykorzystywany do nauki sieci, drugi – do sprawdzania postępów trenowania sieci i zakończenia procesu uczenia przed wystąpieniem tzw. przeuczenia sieci – zjawiska polegającego na nadmiernym dopasowaniu do konkretnego zbioru przypadków kosztem utraty zdolności uogólniania wiedzy. Często wydziela się ze zbioru danych także podzbiór danych walidacyjnych, który pozwala oceniać jakość sieci i porównywać różne warianty sieci zastosowanych do konkretnego zadania.



Rys. 1.5 Przebieg uczenia sieci neuronowej

W trakcie trwania procesu uczenia sieci obserwuje się zmniejszanie wartości błędu (sumy kwadratów różnic pomiędzy zarejestrowanymi i obliczonymi przez sieć wartościami wyjściowymi), zarówno dla próby uczącej jak i testowej. Uczenie jest przerywane, gdy stwierdzone zostanie zwiększenie się błędu obliczonego dla próby testowej. Przykładowy przebieg procesu uczenia sieci pokazany został na rys. 1.5. Trenowanie sieci zostało zakończone po 30 cyklu uczenia. Wartość błędu na osi rzędnych na rys. 1.5 została unormowana do 1.

1.4 WPŁYW ARCHITEKTURY SIECI NEURONOWEJ NA DOKŁADNOŚĆ OPISU ZMIAN REJESTROWANEJ SEJSMICZNOŚCI INDUKOWANEJ

Przystępując do rozwiązywania problemów z wykorzystaniem sieci neuronowych konieczne jest ustalenie typu oraz architektury zastosowanej sieci. Niestety, nie jest znana ogólna metoda projektowania optymalnej architektury sieci. Przykłady proponowanych rozwiązań tego problemu można znaleźć w pracy [4].

W zagadnieniach aproksymacji funkcji najczęściej wykorzystywane są sieci jednokierunkowe, wielowarstwowe, składające się z warstwy wejściowej, jednej bądź większej liczby warstw ukrytych oraz warstwy wyjściowej. W wielu przypadkach wystarczająca jest sieć składająca się z jednej warstwy ukrytej – taką zastosowano w prowadzonych badaniach. W kolejnym etapie projektowania sieci należy ustalić liczbę neuronów w poszczególnych warstwach oraz ich charakterystyki – funkcje aktywacji.

W celu zbadania wpływu architektury sieci na dokładność opisu zmian wydatku energetycznego wstrząsów przeprowadzono wielowariantowe obliczenia, w trakcie których zmieniano liczbę neuronów w warstwie ukrytej – od 1 do 16 (rys. 1.3). Wykorzystywano trzy postacie funkcji aktywacji: liniową, tangens hiperboliczny oraz logistyczną. Obliczenia wykonano pakietem Statistica [5].

Korzystając z danych o czasoprzestrzennym przebiegu eksploatacji górniczej w rozpatrywanej partii złoża wykonano obliczenia zmian energii właściwej odkształcenia sprężystego w czterech potencjalnie wstrząsogennych warstwach skalnych w siatce regularnie rozmieszczonych (co 25 m) 1189 punktów obliczeniowych. Uwzględniono 7-dniowe jednostkowe przedziały czasu – kroki postępu frontów eksploatacyjnych.

Wartości wskaźników Wsk. 1 ÷ Wsk. 4 obliczano z następujących zależności:

$$\text{Wsk. 1}_{kl} = \sum_{i=1}^{ldt} \Delta\Phi_{iobl(W<1)}^+ \quad \text{– obliczone w punkcie } l \text{ dla } k\text{-tej warstwy skalnej sumy}$$

chwilowych przyrostów energii właściwej, w sytuacji gdy wskaźnik wyężenia $W < 1$,

$$\text{Wsk. 2}_{kl} = \sum_{i=1}^{ldt} \Delta\Phi_{iobl(W<1)}^- \quad \text{– obliczone w punkcie } l \text{ dla } k\text{-tej warstwy skalnej sumy}$$

chwilowych spadków energii właściwej, w sytuacji gdy wskaźnik wyężenia $W < 1$,

$Wsk. 3_{kl} = \sum_{i=1}^{l_{dt}} \Delta\Phi_{i_{obl}(W \geq 1)}^+$ – obliczone w punkcie 1 dla k -tej warstwy skalnej sumy chwilowych przyrostów energii właściwej, w sytuacji gdy wskaźnik wyężenia $W \geq 1$,

$Wsk. 4_{kl} = \sum_{i=1}^{l_{dt}} \Delta\Phi_{i_{obl}(W \geq 1)}^-$ – obliczone w punkcie 1 dla k -tej warstwy skalnej sumy chwilowych spadków energii właściwej, w sytuacji gdy wskaźnik wyężenia $W \geq 1$,

gdzie:

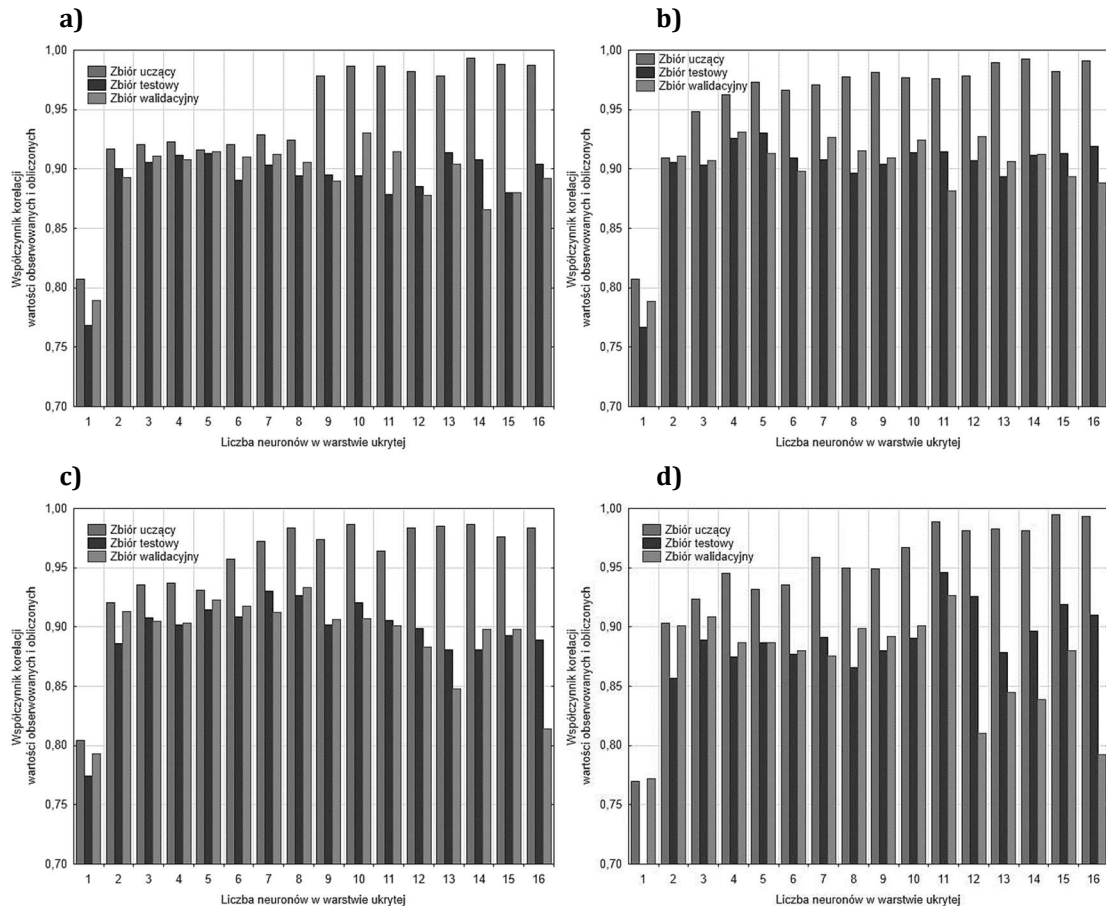
l_{dt} – liczba jednostkowych przedziałów czasu $\Delta t = 7$ dni.

Okres prowadzenia robót podzielono na dwa podokresy:

- 04.2009÷12.2012 – wykorzystując dane o zarejestrowanej sejsmiczności i obliczone wartości wskaźników charakteryzujących zmiany energetyczne w tym przedziale czasu (odpowiadającemu okresowi prowadzenia ściany 002), przeprowadzono naukę sieci neuronowych, różniących się liczbą neuronów oraz ich charakterystykami. Wyniki analiz pozwoliły wytypować optymalną architekturę sieci neuronowej.
- 05.2014÷05.2015 – dla którego, wykorzystując wytrenowaną z wykorzystaniem danych z pierwszego okresu sieć neuronową, przeprowadzono obliczenia zmian gęstości energii wstrząsów. Prognozowane zmiany sejsmiczności porównano z zaobserwowanymi w trakcie prowadzenia robót w polu ściany 003.

Przed przystąpieniem do obliczeń dane charakteryzujące zmiany wartości wskaźników energetycznych oraz wartości gęstości energii wstrząsów zostały podzielone na 3 podzbiory: danych uczących (70% przypadków), danych testowych (15% przypadków) oraz danych walidacyjnych (15% przypadków).

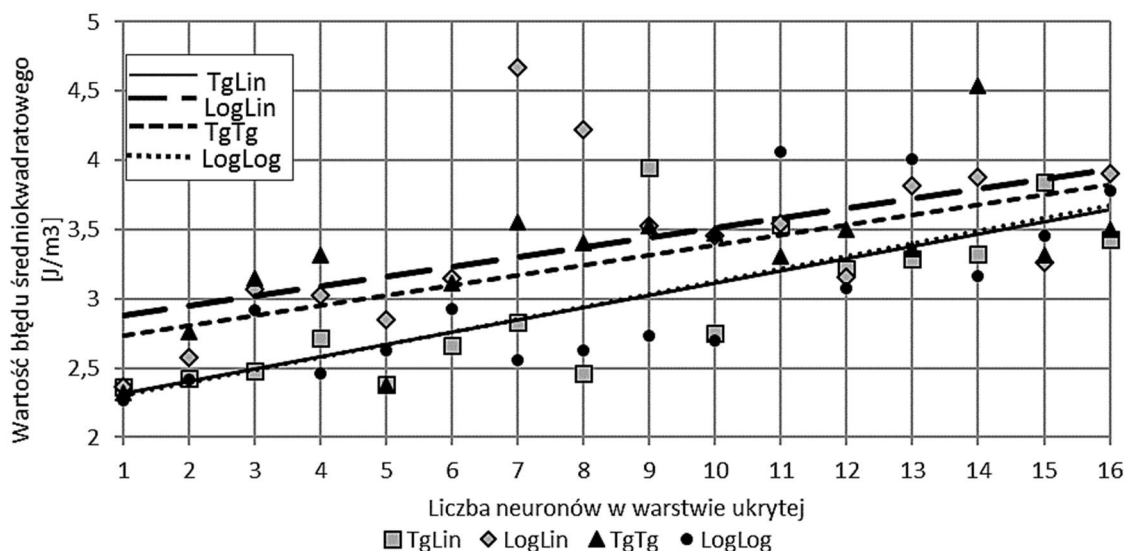
Wyniki przeprowadzonych testów przedstawiono na wykresach zamieszczonych na rys. 1.6. Wykresy prezentują zmiany współczynnika korelacji pomiędzy wartościami otrzymanymi na wyjściu rozpatrywanej sieci i zaobserwowanymi wartościami gęstości energii wstrząsów, osobno dla zbiorów: uczącego, testowego i walidacyjnego. Jak wynika z przeprowadzonych obliczeń, niezależnie od przyjętej charakterystyki neuronów, sieć zawierająca tylko jeden neuron w warstwie ukrytej nie umożliwia dokładnego odtworzenia zaobserwowanych wartości gęstości energii wstrząsów. Zwiększaniu liczby neuronów w warstwie ukrytej sieci towarzyszy wzrost współczynnika korelacji pomiędzy wartościami obserwowanymi i obliczanymi w odniesieniu do zbioru danych uczących. Jednak w miarę wzrostu liczby powyższych neuronów generalnie nie obserwuje się zwiększenia współczynnika korelacji pomiędzy wartościami obserwowanymi i obliczonymi, biorąc pod uwagę zbiory danych testowych i walidacyjnych.



Rys. 1.6 Zmiany współczynnika korelacji pomiędzy wartościami obliczonymi oraz obserwowanymi gęstości energii wstrząsów w zależności od liczby neuronów w warstwie ukrytej i charakterystyki wykorzystanych neuronów w warstwie ukrytej i wyjściowej:
 a) tangens hiperboliczny i liniowa,
 b) logistyczna i liniowa, c) tangens hiperboliczny i tangens hiperboliczny,
 d) logistyczna i logistyczna.

Biorąc pod uwagę ten wniosek wynikający z przeprowadzonych obliczeń, do prognozy zmian wielkości wydatku energetycznego wstrząsów zasadne jest wykorzystywanie sieci zawierającej dwa neurony w warstwie ukrytej, przy czym za wystarczającą należy uznać sieć zawierającą jedną warstwę ukrytą. Dla wszystkich wariantów sieci, różniących się rozpatrywanymi charakterystykami neuronów uzyskano porównywalne dokładności odtworzenia zaobserwowanych zmian gęstości energii wstrząsów. Analizując wykresy na rys. 1.6 można jednak zauważyć bardziej stabilne (biorąc pod uwagę zmiany wartości współczynnika korelacji) zachowanie sieci zbudowanej z neuronów o charakterystyce bipolarnej oraz bipolarnej i liniowej (rys. 1.6a i 1.6c) w porównaniu do sieci zbudowanej z neuronów o charakterystyce unipolarnej oraz unipolarnej i liniowej (rys. 1.6b i 1.6d). W celu dokonania wyboru optymalnej do postawionego zadania charakterystyki neuronów, dla rozpatrywanych sieci, wytrenowanych z wykorzystaniem danych zaobserwowanych podczas prowadzenia ściany 002, wykonana została prognoza wydatku energetycznego wstrząsów obejmująca okres prowadzenia ściany 003. Analiza uzyskanych wyników została przedstawiona na rys. 1.7. Wykres prezentuje zmiany wartości błędu

średniokwadratowego prognozy wydatku energetycznego wstrząsów określone dla wszystkich analizowanych sieci. Na wykresie naniesiono także linie trendu wartości błędów.



Rys. 1.7 Zmiany wartości błędu średniokwadratowego prognozy gęstości energii wstrząsów dla ściany 003 w zależności od liczby neuronów w warstwie ukrytej i charakterystyki wykorzystanych neuronów w warstwie ukrytej i wyjściowej:
 TgLin - tangens hiperboliczny i liniowa, LogLin - logistyczna i liniowa,
 Tg - Tg tangens hiperboliczny i tangens hiperboliczny, Log - Log logistyczna i logistyczna.

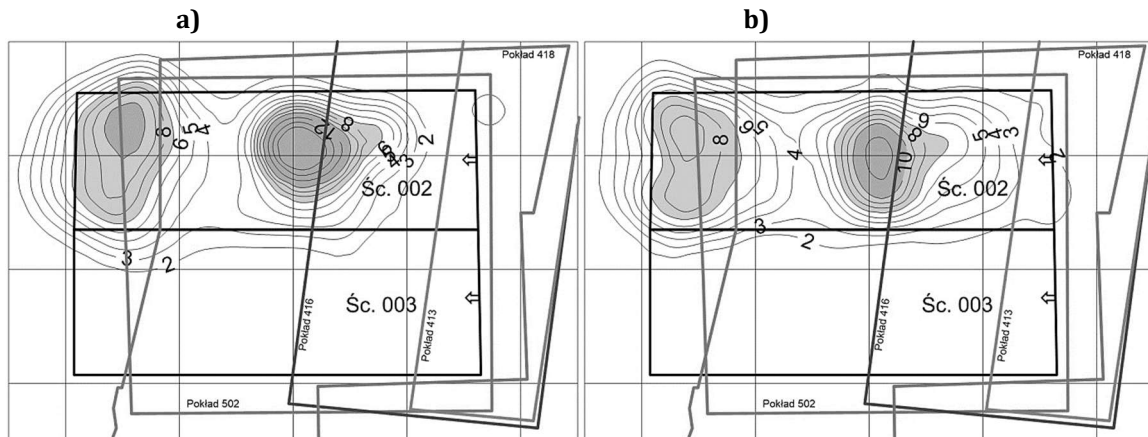
Wyniki prognozy wskazują, że w przypadku sieci zawierającej dwa neurony w warstwie ukrytej, zastosowanie neuronów ukrytych o charakterystyce bipolarnej (tangens hiperboliczny) i wyjściowego o charakterystyce liniowej, pozwala uzyskać dokładniejszą prognozę w porównaniu do sieci wykorzystującej w warstwie ukrytej neurony o charakterystyce unipolarnej.

Przeprowadzone obliczenia testowe wskazują, że do opisu zaobserwowanych i prognozowania zmian gęstości energii wstrząsów najkorzystniej jest zastosować sieć zawierająca dwa neurony w pojedynczej warstwie ukrytej, przy przyjęciu tangensa hiperbolicznego jako funkcji aktywacji neuronów ukrytych, natomiast funkcji liniowej w przypadku neuronu wyjściowego. Oczywiście wnioski te dotyczą danych pochodzących z rejonu ścian 002 i 003. Dalsze obliczenia, wykonane dla innych rejonów, pozwolą ocenić ogólność przedstawionych wniosków.

Tak prosta budowa sieci, która w pewnych przypadkach może powodować stosunkowo słabe odwzorowanie zbioru danych uczących, zabezpiecza bardzo skutecznie przed zbytnim dopasowaniem do tego zbioru, kosztem utraty informacji o charakterystycznych związkach pomiędzy danymi wejściowymi a sygnałem wyjściowym. Wystąpienie takiego zjawiska (tzw. przeuczenie sieci) praktycznie uniemożliwia wykorzystanie danej sieci w celach prognostycznych.

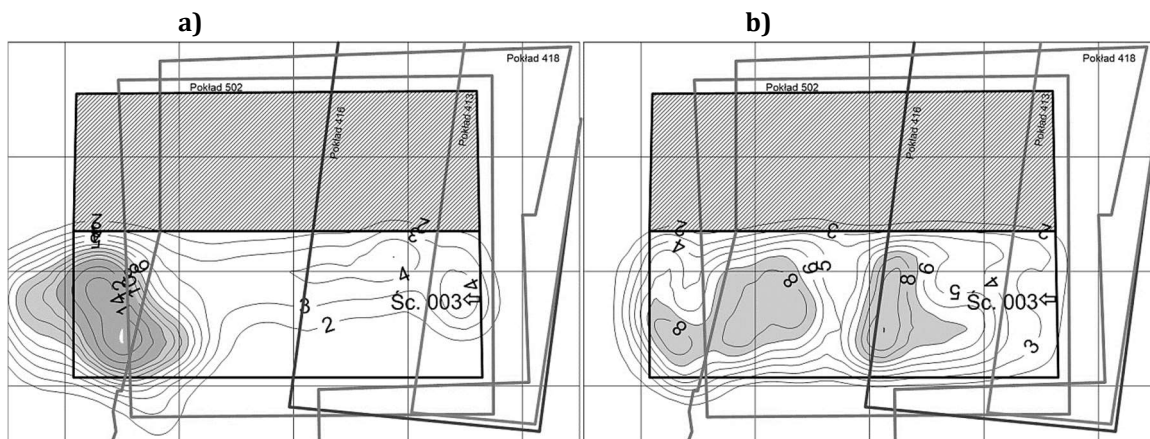
Dla rozpatrywanego w artykule przypadku robót górniczych, pomimo prostoty zastosowanej sieci, udało się uzyskać dokładny opis zaobserwowanych w trakcie

przewodzenia ściany 002 zmian gęstości energii sejsmicznej. Na rys. 1.8 zaprezentowano zaobserwowane (a) i obliczone (b) siecią neuronową rozkłady gęstości energii wstrząsów. Wyznaczona wartość współczynnika korelacji liniowej pomiędzy rozpatrywanymi wielkościami wynosi 0,91. Wartość błędu średniokwadratowego wynosi $1,28 \text{ J/m}^3$.



Rys. 1.8 Rozkład wartości zaobserwowanej (a) i obliczonej siecią neuronową (b) gęstości energii wstrząsów [J/m^3] w okresie prowadzenia ściany 002

Jak już wcześniej wspomniano, sieć neuronowa wytrenowana w oparciu o dane dotyczące sejsmiczności indukowanej w trakcie prowadzenia ściany 002 została zastosowana do sporządzenia prognozy zmian gęstości energii wstrząsów na wybiegu ściany 003. Wyniki tak sporządzonej prognozy wartości wydatku energetycznego wstrząsów (rys. 1.9b) porównano z wartościami zaobserwowanymi (rys. 1.9a).



Rys. 1.9 Rozkład wartości zaobserwowanej (a) i prognozowanej siecią neuronową (b) gęstości energii wstrząsów [J/m^3] w okresie prowadzenia ściany 003

Również w tym przypadku, zastosowana sieć neuronowa umożliwiła stosunkowo dokładną prognozę wydatku energetycznego wstrząsów na wybiegu ściany 003 (rys. 1.9). Największe niedoszacowanie prognozowanej wartości gęstości energii wstrząsów wynosi $9,8 \text{ J/m}^3$, podczas gdy maksymalne przeszacowanie osiąga mniejszą wartość, wynoszącą $8,3 \text{ J/m}^3$. Wartość błędu średniokwadratowego wynosi około

2,42 J/m³, wartość współczynnika korelacji liniowej pomiędzy wartościami obserwowanymi i prognozowanymi wynosi 0,63. W szczególności, trafnie udało się wskazać silnie zagrożony rejon, którym był końcowy odcinek biegu ściany. Prognoza dla przeważającej części wybiegu ściany 003 była nieco zawyżona. Fakt ten można tłumaczyć tym, że nauczona w oparciu o wyniki obserwacji prowadzonych podczas biegu ściany 002 spodziewała się wzrostu sejsmiczności w rejonie oddziaływania krawędzi wytworzonej w pokładzie 416 (por. rys. 1.8). Ten zwiększony poziom gęstości energii wstrząsów jednak nie wystąpił. Tłumaczyć to można lokalnie korzystniejszymi warunkami geologicznymi, a także bardzo szerokim zakresem rzetelnie wykonywanej profilaktyki tapaniowej. Już przed uruchomieniem, a także podczas biegu ściany 003 prowadzono profilaktykę aktywną w postaci strzelań torpedujących w skałach stropowych (z przecinki ścianowej, z chodników przyścianowych) oraz strzelań wstrząsowych wykonywanych z frontu ściany. Powyższe czynniki nie były uwzględniane w trakcie trenowania sieci i stosowania jej do prognozy wydatku energetycznego wstrząsów.

1.4 PODSUMOWANIE

Obliczenia przeprowadzone dla wybranego rejonu pozwoliły stwierdzić możliwość wykorzystania sieci neuronowej o stosunkowo prostej budowie do opisu zmian gęstości energii wstrząsów w oparciu o szacowane zmiany stanów energetycznych zachodzące we wstrząsogennych warstwach skalnych. Uzyskano stosunkowo dużą zgodność opisu zarejestrowanych zmian – współczynnik korelacji liniowej pomiędzy wartościami obliczonymi i zaobserwowanymi gęstości energii wstrząsów wyniósł 0,91. Dla rozpatrywanego zbioru danych sieć neuronowa o optymalnej architekturze zbudowana jest z szesnastu neuronów w warstwie wejściowej, dwóch neuronów o charakterystyce bipolarnej w warstwie ukrytej oraz jednego neuronu o charakterystyce liniowej w warstwie wyjściowej. Odpowiednio wytrenowana sieć neuronowa została z powodzeniem wykorzystana do oceny poziomu sejsmiczności na wybiegu projektowanej kolejnej ściany. Wartość współczynnika korelacji pomiędzy wartościami obliczonymi i zaobserwowanymi gęstości energii wstrząsów wyniosła 0,63.

LITERATURA

1. P. Bańka, T. Cichy, P. Kołodziejczyk. „Wykorzystanie sieci neuronowych do prognozy zagrożenia sejsmicznego.” *Przegląd Górniczy*, nr 4, pp. 1-6, 2016.
2. H. Gil. *The Theory of Strata Mechanics*. Warszawa: PWN, 1991.
3. S. Ossowski. *Sieci neuronowe w ujęciu algorytmicznym*. Warszawa: WNT, 1996.
4. L. Rutkowski. *Metody i techniki sztucznej inteligencji*. Warszawa: PWN, 2012.
5. StatSoft, Inc. (2011). STATISTICA (data analysis software system), version 10.

Data przesłania artykułu do Redakcji: 10.2016

Data akceptacji artykułu przez Redakcję: 03.2017

dr hab. inż. Piotr Bańka, prof. Pol. Śl.
Politechnika Śląska,
Wydział Górnictwa i Geologii
Instytut Eksploatacji Złóż
ul. Akademicka 2, 44-100 Gliwice, Polska
e-mail: piotr.banka@polsl.pl

mgr inż. Tomasz Cichy
Politechnika Śląska,
Wydział Górnictwa i Geologii
Instytut Eksploatacji Złóż
ul. Akademicka 2, 44-100 Gliwice, Polska
e-mail: tomasz.cichy@polsl.pl

DOBÓR ARCHITEKTURY SIECI NEURONOWEJ WYKORZYSTYWANEJ DO OPISU ZMIAN WIELKOŚCI SEJSMICZNOŚCI INDUKOWANEJ

Streszczenie: W artykule przedstawiono wyniki badań mających na celu ustalenie optymalnej architektury sieci neuronowej wykorzystywanej do prognozy zmian gęstości energii wstrząsów indukowanych prowadzonymi robotami górniczymi. Zbiór wykorzystywany do nauki sieci neuronowej zawierał wartości wskaźników charakteryzujących zmiany energii właściwej odkształcenia sprężystego (szacowane z wykorzystaniem rozwiązania przemieszczeniowego zadania brzegowego przestrzennej teorii sprężystości podanego przez H. Gila), a także informacje o zarejestrowanym poziomie sejsmiczności, charakteryzowanym przez określenie gęstości energii wstrząsów. W trakcie prowadzonych badań przeanalizowano wpływ liczby neuronów w warstwie ukrytej sieci (od 1 do 16) oraz przyjętej charakterystyki neuronu (rozważano następujące funkcje aktywacji: liniową, logistyczną i tangens hiperboliczny) na dokładność opisu zmian sejsmiczności indukowanej przy zastosowaniu odpowiednio wytrenowanej (nauczonej) sieci neuronowej. Sprawdzone także wpływ architektury sieci neuronowej na dokładność prognozy zmian wydatku energetycznego wstrząsów górotworu. Rezultaty obliczeń, wykonanych dla silnie zagrożonego wstrząsami i tąpniętami rejonu robót górniczych prowadzonych w jednej z kopalń węgla kamiennego GZW, potwierdziły możliwość zastosowania sieci neuronowych do szacowania zmian gęstości energii wstrząsów towarzyszących eksploatacji złoża. Stwierdzono także, że wystarczającą dokładność prognozy można uzyskać stosując mało rozbudowaną sieć neuronową, zawierającą w warstwie ukrytej tylko dwa neurony, natomiast jako funkcję aktywacji korzystnie jest przyjąć tangens hiperboliczny.

Słowa kluczowe: energia odkształcenia sprężystego, sejsmiczność indukowana, sieci neuronowe

**THE ARCHITECTURE CHOICE OF THE NEURAL NETWORK USED
TO DESCRIBE CHANGES IN SIZE OF INDUCED SEISMICITY**

Abstract : *The article presents results of researches which have been done to define optimal architecture of neural network used to forecast changes in size of seismicity induced by mining works in progress. The set that had been used to learn neural network has included value of indicators that characterize changes of elastic strain energy (estimated with edge task translocation of spatial theory of elasticity given by H. Gil) and information about recorded level of seismicity, described by density of seismic events energy. While doing the research the influence of neurons number in the hidden layer of the network (from 1 to 16) was analyzed and neuron's characteristic (following activation functions have been considered: linear, logistic and hyperbolic tangent) for accuracy in describing induced seismicity changes when appropriately trained neural network is applied. Influence of neural network architecture on exactitude of change forecast of energy expenditure of rock mass shocks has also been checked. Calculation results for the region that is extremely at risk of shocks (region of mining works that are in progress in one of GZW coal mines) have confirmed possibility of using neural networks to estimate changes in size of induced seismicity connected with deposit exploitation. It has been also stated that forecast accuracy can be achieved by using not a very expanded neural network consisted of only two neurons in the hidden layer, however as an activation function it is favorably to use hyperbolic tangent.*

Key words: *elastic strain energy, induced seismicity, neural networks*