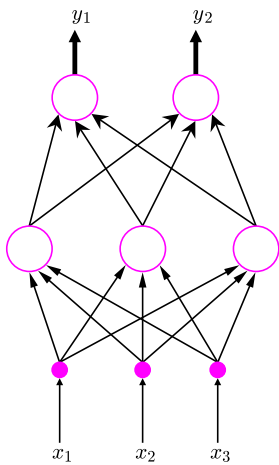


Rys.1. Schemat sztucznego neuronu.



Rys.2. Prosta sieć neuronowa.

Już w pierwszej połowie lat 40. XX wieku, w czasach gdy cybernetyka dopiero zaczynała zdobywać popularność, a wszystkie elektroniczne maszyny arytmetyczne (zwane dziś komputerami) na świecie można było policzyć na palcach jednej ręki, badacze zaczęli się zastanawiać nad mechanizmami, które leżą u podstaw zdolności ludzi i zwierząt do wykonywania skomplikowanych zadań. Niezwykłą własnością mózgu u ludzi i zwierząt jest zdolność do przeprowadzania bardzo zaawansowanych operacji w oparciu o częstokroć zaburzoną, niekompletną i nieściśłą informację. Jednocześnie, nasz układ nerwowy składa się z względnie prostych elementów zwanych neuronami, które w porównaniu z wysokopoziomowymi zdolnościami abstrakcyjnego myślenia wydają się wręcz trywialne. Zatem zdolności człowieka do, na przykład, rozpoznawania pisma ręcznego nie biorą się z budowy samej komórki nerwowej, lecz ze sposobu, w jaki miliardy takich komórek łączą się i współpracują.

Pierwsi badacze w dziedzinie sztucznych sieci neuronowych zadawali sobie jasno sprawę z tego, że sam neuron nie stanowi o myśleniu. Jednak, aby wykonać pierwszy krok, skoncentrowali się na skonstruowaniu elementu podstawowego – sztucznego neuronu. Sztuczny neuron jest daleko idącym uproszczeniem prawdziwej komórki nerwowej. W podstawowym modelu, który do dziś niewiele się zmienił, sztuczny neuron odbiera sygnały wejściowe x_1, \dots, x_n ze środowiska lub od innych neuronów. Sygnały wejściowe, zwykle liczby rzeczywiste, są następnie mnożone przez przypisane im wagi w_1, \dots, w_n . Wagi odpowiadają za to, jaki jest wpływ konkretnego sygnału x_i na działanie neuronu. Działanie neuronu polega na zsumowaniu wartości sygnałów wejściowych pomnożonych przez wagi i podaniu ich jako argumentu do *funkcji aktywacji* f . Wartość funkcji aktywacji $y = f(\sum_{i=1}^n w_i x_i)$ stanowi odpowiedź sztucznego neuronu i jest zwracana do środowiska lub przesyłana do innych neuronów (patrz rysunek 1). „Programowanie” neuronu polega na takim dobraniu wag, aby neuron podawał oczekiwaną odpowiedź na zestaw sygnałów wejściowych. W przypadku pojedynczego neuronu taką technikę stosunkowo łatwo skonstruować, gorzej gdy trzeba poustawiać wagi całej sieci neuronów (na przykład takiej jak na rysunku 2).

Brak dobrego sposobu na ustalanie wag w bardziej złożonych, wielowarstwowych strukturach sztucznych sieci neuronowych stanowił przez długi czas duży problem dla badaczy w tej dziedzinie. Zagadnienie to było niezwykle ważne, gdyż już pod koniec lat 60. XX wieku wiadomo było, że pojedyncze neurony mają bardzo ograniczone praktyczne zastosowania. Brak dobrych metod uczenia złożonych sieci doprowadził na przełomie lat 60. i 70. do znacznego ograniczenia funduszy na badania w tej dziedzinie. Dopiero w latach 80. rozpowszechniła się metoda znana jako *wsteczna propagacja błędu* (*ang. error backpropagation*), która pozwoliła na efektywne *uczenie* sieci właściwego układu wag. Podstawą metody wstecznej propagacji są własności arytmetyczne pochodnych pewnych szczególnych wielowymiarowych funkcji złożonych. Własności te, scharakteryzowane w połowie lat 70. przez P. Werbosa w jego pracy doktorskiej, pozwoliły na opracowanie efektywnej procedury doboru wag sieci neuronowych do stawianych wymagań.

Dzisiaj, po ponad 50 latach od pierwszych prac, sztuczne sieci neuronowe zajmują mocną pozycję w wielu działach nauki i techniki. Wiele nowoczesnych urządzeń codziennego użytku, szczególnie tych rodem z Kraju Kwitnącej Wiśni, wykorzystuje elementy sterujące oparte na sieciach neuronowych. Zaletą tych rozwiązań jest elastyczność i odporność na zakłócenia.

Jednym z najlepiej znanych i spektakularnych zastosowań sztucznych sieci neuronowych jest sortowanie listów. W wielkich sortowniach przesyłek w USA i Kanadzie stosuje się maszyny, które rozpoznają napisany na kopercie (często pismem odręcznym) kod pocztowy i następnie drukują na niej kod kreskowy za pomocą którego inne maszyny wykonują sortowanie. Pozwala to zidentyfikować i odczytać 5 do 8 kodów na sekundę z dokładnością przekraczającą 90%. Przy milionach przesyłek przechodzących codziennie przez takie centra pocztowe jest to bardzo znaczące udogodnienie.

*Instytut Matematyki Uniwersytetu Warszawskiego