

„Im więcej warstw, tym większa szansa,
że nikt nie wie, co się tam dzieje.”

*Sztuczne sieci neuronowe są sposobem
na modelowanie złożonych wzorców,
podobnie jak ludzki mózg.*
Geoffrey Hinton

4.1. Geneza i rozwój sztucznych sieci neuronowych

Historia sztucznych sieci neuronowych (ANN – Artificial Neural Network) daje początek zaawansowanym narzędziom inteligencji obliczeniowej (*Soft Computing*), które po dzień dzisiejszy budzą wielkie nadzieje i liczne kontrowersje. Naukowe zainteresowanie modelami biologicznych sieci neuronowych jest nawet starsze od idei komputera von Neumanna (1944) i datuje się na wczesne lata 40. XX wieku. Pionierami tych badań byli McCulloch i Pitts, którzy w roku 1943 przedstawili pierwszy formalny model neuronu (perceptronu) w postaci układu arytmetyczno-logicznego, którego idea do dziś stanowi przedmiot intensywnych badań i wielu oryginalnych aplikacji.

Koncepcja sztucznych sieci neuronowych pochodzi od słynnego perceptronu F. Rosenblatta i Ch. Wightmana. Zaproponowane przez nich w latach 60. XX w. jako cybernetyczne symulatory komórki nerwowej. Takie podejście do neuronu okazało się bardzo atrakcyjne poznawczo i było intensywnie rozwijane przez wielu badaczy w różnych placówkach naukowych. Sztuczne sieci neuronowe sukcesywnie rozwijane od pojedynczej komórki poprzez struktury jednowarstwowe, aż do modeli wielowarstwowych oferowały nieznaną dotychczas zdolność programowego uczenia i maszynowego samodoskonalenia.

W latach 80. XX wieku opublikowano szereg prac, które w istotny sposób wpłynęły na rozwój badań nad sztucznymi sieciami neuronowymi, a przełomowym wydarzeniem było zaproponowanie w roku 1982 przez J. J. Hopfielda – laureata nagrody Nobla w dziedzinie fizyki nowej, rekurencyjnej architektury pamięci sieci neuronowych, opartej na idei sprzężenia zwrotnego. Dekada lat 80. to także rewolucja mikroelektroniczna i pojawienie się najpierw mikrokomputerów, a potem za sprawą IBM słynnych komputerów osobistych PC (Personal Computer). W efekcie po raz pierwszy w dziejach ludzkości komputery trafiły „pod strzechy” i stały się nie tylko kultowym przedmiotem pożądania, ale zwykłym narzędziem pracy, a przede wszystkim wspaniałym gadżetem naukowym, domowym i osobistym imitatorem sztucznego intelektu.

Jedną z najważniejszych zalet sieci neuronowych ANN był fakt, że do pracy nie wymagały one specjalnego oprogramowania, a ich działania opierały się bardziej na numerycznych procedurach niż na zaawansowanych programach komputerowych. Wystarczyło bowiem raz zbudować taką sieć logiczną, a ona

sama była w stanie nauczyć się swojej funkcji, którą mogła wykonywać dowolnie długo. Rola programisty polegała na zaprojektowaniu struktury sieci, która najlepiej będzie spełniała założenia warunkujące rozwiązanie danego problemu. Jeszcze inną zaletą sztucznych sieci ANN jest fakt polegający na tym, że w zwykłym programie nawet najmniejszy błąd jest powodem wadliwego działania, a nawet zawieszenia pracy programu. Sieć neuronowa sama potrafiła do pewnego stopnia omijać takie pułapki i dopiero po przekroczeniu pewnego progu błędów mogą wystąpić problemy z jej funkcjonowaniem. Zasadniczym walorem sieci neuronowych na tym etapie okazała się stosunkowo prosta ich implementacja programowa (komputerowa) wynikająca z braku głębszych podstaw metodologicznych i stosunkowo płytkiej warstwy teoretycznej. Do zbudowania „prywatnej” sieci neuronowej wystarczała znajomość mechanizmu jej funkcjonowania i dostęp do średniozaawansowanych narzędzi i aplikacji programowych.

Podstawą dalszych rozważań będzie opisowa definicja zaproponowana przez R. Tadeusiewicza (1993), który sztuczną sieć neuronową określa wprost jako „bardzo uproszczony model mózgu”. Składa się ona z dużej liczby (od kilkuset do kilkudziesięciu tysięcy) elementów logicznych (programowych) przetwarzających informację. Elementy te nazywane są neuronami, chociaż w stosunku do rzeczywistych komórek nerwowych ich funkcje są bardzo uproszczone, by nie powiedzieć – prymitywizowane. Neurony są powiązane w sieć za pomocą połączeń o parametrach wagowych, modyfikowanych w trakcie tzw. procesu uczenia maszynowego. Topologia połączeń oraz ich parametry stanowią program działania sieci, zaś sygnały pojawiające się na jej wyjściach, w odpowiedzi na określone sygnały wejściowe, są rozwiązaniami stawianych jej zadań.

4.2. Idea działania sztucznych sieci neuronowych

Impulsem do podjęcia badań nad sztucznymi sieciami neuronowymi była modelowa koncepcja funkcjonowania ludzkiego mózgu. Jak wiadomo mózg ludzki składa się z wielkiej liczby wyspecjalizowanych komórek nerwowych zwanych neuronami, których najistotniejszą cechą jest zdolność do autonomicznego przetwarzania i wielokierunkowego przekazywania informacji między sobą. Neurony posiadają typową budowę, jak każda żywa komórka – jądro, sigma ciało, błona komórkowa, a dodatkowo są wyposażone w specjalne wypustki dendryty, z których jeden wyróżniony tzw. neuryt zwany też aksonem kończy się pajęczyną drobniejszych rozgałęzień zakończonych specjalną szczeliną synaptyczną. Aksony niektórych neuronów są krótkie, ale niektóre mogą być bardzo długie.

Dendryty zbierają sygnały z sąsiednich neuronów i przekazują je do jądra macierzystej komórki nerwowej – neuronu, który podejmuje decyzje o dalszej transmisji odpowiednio przetworzonego sygnału na zewnątrz. Gdy całkowite natężenie zebranych sygnałów wejściowych przekroczy pewien próg, zwany potencjałem aktywacji, wówczas za pośrednictwem aksonu jest wysyłany na zewnątrz określony impuls elektryczno-chemiczny. Jeśli sygnał dotrze do końca aksonu, może zostać przejęty za pomocą odpowiedniej synapsy przez sąsiednie neurony znajdujące się na jego zakończeniu. Przeskok sygnałów do kolejnego neuronu nie następuje bezpośrednio lecz za pośrednictwem specjalnej szczeliny zwanej synapsą, która przekazuje albo impulsy elektryczne, albo specjalne substancje chemiczne zwane neurotransmiterami. Potencjał aktywacji neuronu uruchamia końcówkę aksonu celem przesłania neurotransmitera do synapsy i dalszego przekazania tego sygnału do odpowiedniego neuronu.

Sygnały przechodzące przez synapsę podlegają dodatkowej transformacji i generalnie mogą być wzmocnione lub osłabione w zależności od stanu synapsy. W sensie formalnym można powiedzieć, że synapsa dysponuje pewnym współczynnikiem wagowym, który rzutuje na ostateczny kształt sygnału wyjściowego. Ta właściwość (multiplikatywna lub addytywna) synapsy ma ogromny wpływ na sposób w jaki mózg przetwarza informację. Miliardy wag synaptycznych aktywnie wpływają na przemieszczane w mózgu sygnały i prawdopodobnie w ten sposób objawia się kreatywność mózgu. Spostrzeżenia te doprowadziły kanadyjskiego psychologa Donalda Hebba do postawienia hipotezy, że regulowanie wag synaptycznych jest podstawą zapamiętywania i uczenia się człowieka. Hipoteza ta, nosząca nazwę teorii Hebba, została sformułowana w roku 1949 i stała się podwaliną psychologii poznawczej, a obecnie kardynalną zasadą funkcjonowania sieci neuronowych ANN.

Z punktu widzenia sztucznych sieci neuronowych najważniejszymi częściami składowymi neuronu są:

- jądro stanowiące centrum obliczeniowe neuronu, gdzie zachodzą procesy kluczowe dla jego funkcjonowania,
- sieć dendrytów, stanowiących informacyjne wejścia neuronu, którymi do jego wnętrza trafiają sygnały poddane późniejszej obróbce,
- akson (neuryt), jako informacyjne wyjście neuronu za pomocą którego sygnał wędruje do innych neuronów poprzez szczeliny synaptyczne,
- kolekcja szczelin synaptycznych, będąca zakończeniem aksonu, przez które przepływający sygnał wyjściowy może być dodatkowo wzmocniany lub osłabiany.

Reasumując, sztuczne sieci neuronowe powstały dzięki obserwacjom i próbom modelowego naśladowania przetwarzania informacji w centralnym systemie nerwowym organizmów żywych, głównie kręgowców. Są one prostymi modelami naśladowującymi funkcje biologicznego, centralnego systemu nerwowego w zakresie przetwarzania informacji. W systemach sztucznych sieci

neuronowych sposób przetwarzania informacji jest „zakodowany” strukturą sieci i jej parametrami nazywanymi wagami i progami. Ten prosty model ludzkiego mózgu, w zasadzie jego elementarnej komórki nerwowej – neuronu (perceptronu) był podstawą sformułowania koncepcji i budowy sztucznych sieci neuronowych ANN, które na długie dziesiątki lat skierowały naukę na niepewne tory naiwnych analogii mózg – komputer. Aktualnie podejście to dominuje w obszarze sztucznej inteligencji generatywnej, gdyż jest postawą uczenia maszynowego, które notuje wielkie tryumfy w praktycznym wykorzystaniu sztucznej inteligencji, czego najlepszym przykładem jest czat GPT 3.0.

Pomijając kwestie czysto narzędziowe, głównie numeryczne i obliczeniowe zasadniczym mankamentem sztucznych sieci neuronowych, różniących je diametralnie od biologicznych struktur ludzkiego mózgu – przy dzisiejszym stanie wiedzy, jest brak świadomości i własnej tożsamości tych sztucznych produktów nauki i technologii. A bez świadomości nie można mówić o jakiegokolwiek inteligencji, nawet sztucznej, może raczej obliczeniowej (komputerowej). Omawiane sieci neuronowe są projektami symbolicznymi (programy komputerowe) lub materialnymi (chipy elektroniczne), będące wynikiem badań naukowych i sztuki inżynierskiej wzorowanych na genialnych dziełach przyrody, kształtowanych w toku długowiekowej ewolucji gatunków i ludzkiej cywilizacji.

Problematyka badawcza sztucznych sieci neuronowych wywodzi się z rodziny nauk dedukcyjnych, których podstawy teoretyczne opierają się na założeniu, że jeśli dana sieć neuronowa poświęci dostatecznie dużo czasu na rozwiązywanie zadań treningowych, to jej odpowiedzi będą coraz bardziej zbliżone do efektywnego rozwiązania rzeczywistego problemu. Po latach eksperymentów, pozbawionych niestety solidnych podstaw teoretycznych okazało się, że sieci neuronowe rzeczywiście rozwiązują niektóre nawet skomplikowane zadania w sposób zadawalający. Dopiero wdrożenie tzw. generatywnych sztucznych sieci neuronowych GenAI, opartych na statystycznej analizie wielkich zbiorów danych i mechanizmie uczenia maszynowego otworzyło przed nimi bardzo atrakcyjne obszary aplikacyjne, niemal we wszystkich dziedzinach sztucznej inteligencji.

4.3. Przetwarzanie informacji w sieci neuronowej

Centralną pozycję w teorii sztucznych sieci neuronowych zajmuje sformalizowany model neuronu (perceptron), zobrazowany za pomocą elementarnej jednostki logiczno-arytmetycznej przeznaczonej do przetwarzania zewnętrznych sygnałów źródłowych na sygnały wyjściowe, sterujące pracą, np. innych neuronów. Sztuczna sieć neuronowa jest odpowiednio skonfigurowanym układem sztucznych (matematycznych, programowych, a niekiedy także

sprzętowych) neuronów, połączonych między sobą licznymi więzami, poprzez specjalny system „mikroprzełączników” zwanych synapsami.

Wbrew pozorom w strukturze sztucznych sieci neuronowych kluczową rolę pełnią bardziej synapsy niż neurony. Synapsy są najbardziej aktywnymi elementami tych sieci, gdyż poprzez współczynniki wagowe bezpośrednio wpływają na wielkość sygnału przetwarzanego przez neurony. Stosownie do wartości współczynnika wagowego sygnały te na wyjściu synaptycznym mogą być bardzo prosto wzmacniane lub osłabiane. Tak przetworzone sygnały kierowane są do centrum zarządzania, czyli do jądra komórki nerwowej. Sztuczny neuron przejmuje nadchodzące sygnały wejściowe i rozpoczyna ich interpretację – sumowanie za pomocą specjalnych formuł, w tym przypadku prostych funkcji matematycznych, np. sigmoidalnej. Dodatkowo może być aktywowana specjalna funkcja transformacji sygnału wyjściowego do żądanej postaci analitycznej. Najważniejszym odkryciem tego modelu było sumowanie sygnałów wejściowych z odpowiednią wagą i poddanie na wyjściu otrzymanej sumy działaniu nieliniowej funkcji aktywacji. Wynikowy sygnał sumaryczny jest przesyłany za pomocą właściwego aksonu poprzez szczelinę synaptyczną do innych neuronów, zgodnie z aktualną konfiguracją sieci połączeń.

Funkcje, których używa sztuczny neuron do przekształcenia sumy przychodzących sygnałów wejściowych w sygnał wyjściowy nazywamy funkcją aktywacji. W mózgu aktywacja neuronu następuje skokowo po przekroczeniu pewnej granicy sumarycznego natężenia sygnałów przychodzących ze wszystkich neuronów. Na gruncie sztucznych sieci neuronowych ANN modelowanie takiego zachowania oznacza zastosowanie pewnej progowej funkcji matematycznej (aktywacji).

Należy zwrócić uwagę, że faktycznie sztuczna sieć neuronowa nie zawiera żadnych sztucznych (biologicznych) neuronów, a są to najczęściej instrukcje (kody) programu komputerowego, który realizuje logiczną strukturę pewnej sieci organizacyjno-funkcjonalnej. Inaczej można powiedzieć, że sztuczna sieć neuronowa to program komputerowy symulujący strukturę i logikę działania pewnej (biologicznej) sieci neuronowej (nerwowej). Jeszcze inaczej, sztuczna sieć neuronowa to rodzaj wirtualnej sieci, która z różnym skutkiem imituje zachowania swojego naturalnego pierwowzoru, czyli struktur mózgowych. Istnieje ona z reguły w pamięci komputera lub rzadziej może mieć swoją fizyczną reprezentację sprzętową.

Klasyczny model tzw. sieci perceptronowej składa się z trzech warstw zawierających neurony warstwy – wejściowej, pośredniej (ukrytej) i wyjściowej. Neurony warstwy wejściowej służą do wprowadzania impulsów (wymuszeń) zewnętrznych. Neurony warstwy wyjściowej przekazują wyniki obliczeń na zewnątrz sieci. Pomiędzy warstwą wejściową a wyjściową typowa sieć neuronowa ma jedną lub więcej warstw pośrednich (ukrytych), gdzie wykonywana jest większość obliczeń. Liczba elementów w każdej warstwie nie jest niczym ograniczona i może być dowolnie kształtowana. Niektóre sieci

badawcze zawierały ponad 100 neuronów, co oczywiście implikowało znaczną liczbę połączeń, między którymi przepływały różne sygnały. W strukturze prostej sieci perceptronowej nie występują sprzężenia zwrotne.

Największą zaletą aplikacyjną sieci neuronowych jest możliwość ich programowego uczenia i adaptacji do zadanych wzorców. Aby zrealizować to zadanie nie musimy posiadać szczegółowej wiedzy o badanym procesie, a zwłaszcza nie musimy budować żadnego modelu matematycznego, czy programu komputerowego. Zdolność sieci do efektywnego uczenia się, zastępuje konieczność tradycyjnego modelowania i programowania nowych stanów i procesów. Na podstawie sygnałów wejściowych oraz sygnałów zadanych (wzorcowych) sieć neuronowa jest w stanie samoistnie nauczyć się w kolejnych cyklach iteracyjnych, jak danym procesem należy sterować, aby spełnić zakładane wymagania i uzyskać pożądaną wynik.

Wielkim atutem sztucznych sieci neuronowych może być efektywny proces ich uczenia, zastępujący tradycyjne programowanie. W odróżnieniu od komputerów, które są programowane do realizacji określonych zadań, większość sieci neuronowych jest uczona (trenowana), np. na podstawie kojarzenia obrazów i odpowiednich zależności funkcyjnych. Zamiast pisania programów projektanci sieci neuronowych określają najlepszą architekturę sieci, specyfikują charakterystyki neuronów i początkowe wartości wag oraz wybierają metodę uczenia sieci.

4.4. Mechanizm działania sztucznych sieci neuronowych

Modelowy mechanizm działania sztucznej sieci neuronowej (perceptronowej) zostanie wyjaśniony na prostym przykładzie. Na wstępie zakładamy, że każdy neuron otrzymuje sygnały sterownicze od wszystkich neuronów znajdujących się po jego stronie wejściowej. Sygnały te odpowiadają liczbom rzeczywistym, które mogą zmieniać się w sposób ciągły w określonym przedziale wartości. Wobec tego liczby te można traktować jako ciąg impulsów nerwowych o określonej amplitudzie i częstotliwości, choć interpretacja tych liczb w literaturze jest dość niejednoznaczna, co zawsze rzutuje na efektywność budowanego modelu.

Pojedynczy neuron zbiera wszystkie sygnały, które dotarły do niego przez sieć połączeń wejściowych i przykładowo w sposób arytmetyczny sumuje te sygnały. Otrzymany wynik sumaryczny jest następnie przekształcany za pomocą pewnej nieliniowej funkcji, np. sigmoidalnej, dość powszechnie znanej m.in. w pedagogice, psychologii czy w technice. Każdy sygnał wejściowy zostaje więc poddany obróbce matematycznej za pomocą tej funkcji, co oznacza że suma sygnałów wejściowych jest transformowana na postać funkcji sigmoidalnej.

W pedagogice krzywa sigmoidalna reprezentuje klasyczną funkcję uczenia, np. dzieci. Zgodnie ze sprawdzoną teorią na początku dzieci uczą się powoli i z pewnymi trudnościami. Potem następuje okres fascynacji własnymi zdolnościami poznawczymi i efekty są doskonałe. Jednak po pewnym okresie efektywność nauki ulega osłabieniu i zbliża się do pewnego poziomu nasycenia wiedzą i umiejętnościami i następuje pewnego rodzaju nasycenie.

Aby rozwiązać konkretny problem badawczy za pomocą sztucznych sieci neuronowych w najprostszym przypadku należy dla każdego neuronu:

- po pierwsze, zsumować wszystkie sygnały wejściowe, pochodzące np. od neuronów wejściowych i otrzymaną sumę przekształcić za pomocą np. funkcji sigmoidalnej na sygnał wyjściowy z tego neuronu,
- po drugie, przekształcony sigmoidalnie sygnał wyjściowy przesyłany jest do wszystkich neuronów leżących po jego stronie wyjściowej, tym razem jako ich sygnały wejściowe,
- po trzecie, na końcu należy zsumować wszystkie sygnały wyjściowe neuronów z ostatniej warstwy obliczeniowej, a wartość ta będzie „rozwiązaniem” postawionego na wstępie problemu.

W budowanych sztucznych sieciach neuronowych ANN sprawy znacznie się komplikują, jeśli weźmiemy pod uwagę fakt, że na drodze między dwoma neuronami występuje, wzorem biologicznych neuronów, dodatkowa „śluzka”, tzw. synapsa, która może wzmacniać lub osłabiać przechodzące przez to połączenie sygnały. Stopień wzmocnienia lub osłabienia matematycznie wyrażać będzie pewien współczynnik wagowy, charakteryzujący daną synapsę. Oznacza to, że synapsa mnoży przepuszczany sygnał przez pewną wagę, która stanowi element sterujący (uczący) danej sieci.

Wagi poszczególnych synaps, będące zasadniczym elementem sterującym (decyzyjnym) mogą być racjonalnie zmieniane przez specjalną procedurę programową. Kryterium ich zmian wynika ze stopnia doskonałości czy użyteczności generowanego rozwiązania. Jeśli więc aktualne sygnały wejściowe niepełnie spełniają nasze oczekiwania to za pomocą systemu wag synaptycznych możemy w kolejnych krokach iteracyjnych poprzez zmianę ich wartości stopniowo przybliżać się do zakładanego wzorca. Aby ograniczyć i ukierunkować liczbę prowadzonych iteracji wartości wag synaptycznych, najczęściej stosuje się metodę tzw. propagacji wstecznej. Sposób uczenia sieci wielowarstwowych za pomocą algorytmu propagacji wstecznej zaczyna się od losowo dobranych wag, które stopniowo są poprawiane w kolejnych cyklach symulacyjnych.

Końcowym wynikiem nauczania sieci neuronowej jest właściwie dobrana tablica wag synaptycznych, którą uzyskujemy po mozolnym testowaniu sieci dla coraz lepszych wartości wagowych. Wielokrotne iterowanie algorytmu uczenia sieci powinno doprowadzić do zmniejszenia całkowitego błędu sieci aż do pewnego, akceptowanego punktu. W idealnej sytuacji w punkcie tym całkowity

błąd powinien wynosić zero, co oznacza, że sieć produkuje dokładnie pożądane sygnały wyjściowe dla zadanego wzorca zbioru treningowego. Wypracowane wartości wag synaptycznych zarówno przed, jak i po regulacji przechowuje się w pamięci komputera w postaci specjalnych tablic (macierzy) i stanowią one zasadniczy element funkcjonowania odpowiedniego programu komputerowego. Procedura znajdowania wag synaptycznych ma postać stosunkowo prostego algorytmu iteracyjnego, który wymaga efektywnego i mozolnego oprogramowania i testowania. Językiem uwagi tego algorytmu jest wielokryterialna funkcja oceny rozwiązania końcowego, która w bardziej złożonych przypadkach posiada nie jedno, a wiele ekstremów lokalnych i dlatego problem znalezienia ekstremum globalnego jest bardzo złożony.

4.5. Tablica wag synaptycznych

Bezpośrednim efektem pracy wytrenowanych sztucznych sieci neuronowych są tablice wag synaptycznych, które zawierają wartości liczbowe wag (współczynników) przypisanych do poszczególnych synaps dla danej sieci. Należy mieć świadomość, że wartości tych wag pochodzą ze żmudnych, czasami wielogodzinnych obliczeń podczas, których algorytm propagacji wstecznej z mozołem poprawiał kolejne wersje docelowego rozwiązania. Zbieżna korekta prowadzi do takiego ustawienia wszystkich wag synaptycznych, że rozwiązanie końcowe spełnia założone kryterium jakości, w szczególności optymalności.

Metoda ta nazywa się propagacją wsteczną, gdyż najpierw koryguje się wagi prowadzące do warstwy wyjściowej, potem wagi poprzedzające i tak dalej aż do osiągnięcia warstwy wejściowej. Kolejność poprawiania wag jest więc przeciwna do kolejności w jakiej są obliczane sygnały, gdy sieć wykonuje swoje użytkowe zadania. Metoda ta pozwala tak ustawiać wagi w każdym połączeniu, że zmodyfikowana w ten sposób sieć generuje zbiór sygnałów wyjściowych, który jest coraz bliższy rozwiązaniu właściwemu.

Każde kolejne poprawne rozwiązanie danego problemu automatycznie modyfikuje wagi synaptyczne, poprawiając ich dotychczasowe wartości. Jeśli w danym kroku wynik końcowy nie jest wystarczająco dobry, kolejne zastosowanie metody propagacji wstecznej wynik ten po raz kolejny poprawi. Postępowanie takie należy prowadzić iteracyjnie aż do uzyskania wyniku satysfakcjonującego, a aktualne wartości wag synaptycznych stanowią wzorcowe parametry ogólnej procedury rozwiązującej zadania o podobnym charakterze.

Zastanówmy się jaki jest sens logiczny każdej wagi, których zebrane wartości stanowią „skarb” naszej wytrenowanej sieci neuronowej. W ogólności tablica wag synaptycznych ma charakter dwuwymiarowej, kwadratowej macierzy, której kolumny i wiersze zawierają numery modelowanych neuronów w poszczególnych warstwach. Na przecięciu tych współrzędnych – neuron X i Y

znajduje się wartość liczbowa wagi synaptycznej połączenia pomiędzy neuronem X a Y. W rzeczywistości tablica „prawdziwych” wag synaptycznych jest nieskończenie duża i wielowymiarowa macierz względem przestrzeni i czasu.

Zgodnie z ideą sztucznych sieci neuronowych podczas każdej sesji roboczej sieć ta samorzutnie się doskonali, czyli modyfikuje dotychczasowe wartości wag synaptycznych, np. za pomocą metody propagacji wstecznej. Jeśli w wyniku długich treningów uzyskamy zadawalający poziom generowanych rozwiązań przez sieć neuronową, to tablicę tą możemy schować do „supersejfu”, jako rzeczywisty skarb dotyczący określonej dziedziny nauki czy życia praktycznego.

Przykładowo, systematycznie trenowana sieć medyczna może zawierać wszystkie znane objawy i skutki chorób. W efekcie może być komputerowym diagnostą we wszystkich kwestiach zdrowotnych, jakie trapią ludzkość. Tym sposobem cała diagnostyka medyczna według stanu na określony dzień może być zawarta w supermacierzy składającej się z milionów liczb. Rodzi się naturalne pytanie – jak skorzystać intelektualnie z bogactwa zakłętogo w niemianowanych, bezwymiarowych liczbach, które nie mają żadnego kontekstu medycznego, ani merytorycznego.

Wiedza zawarta w takiej gigantycznej tablicy jest ulotna i niedostępna ani dla praktyki, ani dla nauki. Ponadto wiedza, która została przez sieć neuronową nabyta, jest w tej sieci rozproszona, przez co zupełnie nieprzydatna dla obserwatora. Aby skorzystać z tej tablicy wymagana jest identyczna sztuczna sieć neuronowa, jaka funkcjonowała na etapie jej tworzenia. W sensie naukowym tablica wag synaptycznych nie stanowi jakiegokolwiek wartości, gdyż nauka bazuje na dostępnej i uznanej wiedzy, a nie na bezładnej płątaniu niewymiernych liczb, nawet uporządkowanych w strukturze symbolicznej macierzy .

4.6. Klasyczne rodzaje sztucznych sieci neuronowych

Rozwój sztucznych sieci neuronowych następował ewolucyjnie od struktur najprostszych do modeli coraz bardziej skomplikowanych. Najprostszym rodzajem sieci neuronowych jest wielowarstwowy perceptron reprezentowany topologicznie przez graf skierowany bez pętli i nawrotów (McCulloch i Pitts, 1943). Tradycyjny perceptron zbudowany jest z trzech modelowych warstw: wejściowej, wyjściowej i pewnej liczby neuronów ukrytych znajdujących się między tymi warstwami. Każdy neuron w warstwie poprzedniej jest połączony z każdym neuronem w warstwie następnej (nadrzędnej), a połączenia te są jednokierunkowe i zawsze skierowane od warstwy wejściowej do wyjściowej.

Informacja docierająca do neuronów wyjściowych przepływa przez całą sieć tylko jeden raz zanim dotrze do neuronów wyjściowych, a po drodze jest ona przetwarzana według przyjętego algorytmu. Neurony znajdujące się w warstwie

ukrytej wnoszą najbardziej istotny element w całym procesie przetwarzania, a mianowicie – system wag, za pomocą których dokonuje się transformacja sygnałów wejściowych na wyjściowe. Perceptrony należą do sieci jednokierunkowych nie zawierających żadnych zamkniętych pętli, a sygnały przechodzą przez taką sieć tylko jeden raz – od wejścia do wyjścia.

Rozwinięciem tych sieci są sieci rekurencyjne, które zawierają różne pętle i sprzężenia zwrotne, a informacje między neuronami mogą przepływać w dowolnych kierunkach. Najprostszą siecią rekurencyjną jest sieć Elmana, która formalnie jest dwuwarstwowym perceptronem z dodatkowymi neuronami ukrytymi w warstwie kontekstowej. W odróżnieniu od sieci jednokierunkowych, które zawsze produkują ten sam sygnał wyjściowy przy tym samym sygnale wejściowym sygnały wychodzące z sieci rekurencyjnej zależą nie tylko od aktualnego stanu sygnałów wejściowych, ale też od poprzednich, np. historycznych stanów sieci.

Kluczowy problem pamiętania i nauczania podejmują tzw. sieci Hopfielda, które stanowią bardzo zaawansowane narzędzie badawcze. Uczenie sieci neuronowych polega na dopasowywaniu wag ich połączeń, tak aby sieć działała w sposób pożądany, zgodnie z zakładanym celem. Sztuczny proces uczenia sieci przebiega według pewnego algorytmu uczenia, który rekurencyjnie w kolejnych cyklach, w sposób automatyczny poprawia wagi synaptyczne poszczególnych neuronów. Najprostszy algorytm uczenia pochodzi od Johna Hopfielda i generalnie bazuje na koncepcji Hebba opisującej procesy zapamiętywania w ludzkim mózgu.

Sieć Hopfielda jest przykładem sieci warstwowej ze sprzężeniem zwrotnym (tzw. sieć rekurencyjna), gdzie wyjścia poszczególnych neuronów są podawane z odpowiednimi wagami na wejścia każdego z neuronów. W strukturze tej sieci występują sprzężenia zwrotne z wyjścia każdego neuronu na wejścia wszystkich pozostałych. Wyjście *i*-tego neuronu stanowi jednocześnie *i*-te wyjście sieci neuronowej. Oczywiście do każdego neuronu podpięte jest też wejście, którym doprowadzana jest odpowiednia składowa wektora testowego.

Strukturę sieci Hopfielda można opisać stosunkowo prosto, jako układ wielu identycznych elementów połączonych metodą każdy z każdym. Dlatego jest najczęściej rozpatrywana jako struktura jednowarstwowa. W odróżnieniu od sieci warstwowych typu perceptron, sieć Hopfielda jest najczęściej siecią rekurencyjną, gdzie neurony są wielokrotnie pobudzone w jednym cyklu rozpoznawania, co uzyskuje się poprzez pętle sprzężenia zwrotnego. Punktem startowym tego algorytmu jest wprowadzony na początku wzorec, który w kolejnych cyklach treningowych jest albo modyfikowany, albo doskonalony. Istotną nowość wnoszona przez sieć Hopfielda polega na tym, że są w nich sprzężenia zwrotne, których inne sieci nie uwzględniają.

Wielokrotnie iterowany algorytm Hopfielda prowadzi do zmniejszenia całkowitego błędu sieci do pewnego punktu, w którym sieć produkuje niemal dokładnie pożądany obraz w stosunku do zadanego wzorca ze zbioru

treningowego. Odpowiednio wytrenowana (nauczona) sieć Hopfielda potrafi odtwarzać całe obrazy na podstawie sygnałów wejściowych zawierających jedynie obrazy fragmentaryczne. Modele sieci Hopfielda są jednymi z najczęściej omawianych, badanych i praktycznie wykorzystywanych. Zazwyczaj są stosowane do rozpoznawania lub klasyfikacji obrazów, które są reprezentowane w sposób binarny. Warto także zaznaczyć, że sieć Hopfielda jest często podawana jako przykład pamięci skojarzeniowej lub jako układ do rozwiązywania zadań z zakresu optymalizacji.

Na uwagę zasługują również tzw. niedeterministyczne sieci neuronowe, w których wagi i progi aktywacji są losowo zmieniane w oparciu o funkcje gęstości prawdopodobieństwa. Sieci te wykorzystują znane idee, takie jak symulowane wyżarzanie (maszyna Boltzmann'a) czy Bayesowską teorię decyzji.

4.7. Zastosowania sztucznych sieci neuronowych

Sztuczne sieci neuronowe SNN już w początkowym okresie swojej kariery fascynowały szczególnie świat badaczy i naukowców. Z pierwotnych aplikacji sztucznych sieci neuronowych masowo zaczęli korzystać eksperci z dziedziny medycyny oraz biolodzy, których interesowało modelowanie naturalnych sieci neuronowych w funkcjonujących strukturach organizmów żywych.

- Fizycy widzieli analogię pomiędzy strukturami sieci neuronowych, a nieliniowymi układami dynamicznymi. Matematycy i inżynierowie różnych specjalności widzieli ich rozległe obszary aplikacji w obliczeniach optymalizacyjnych. Informatycy mechanizm sztucznych sieci chcieli wykorzystać do tworzenia równoległych struktur obliczeniowych w różnych obszarach sztucznej inteligencji, a także w systemach symulacji różnych procesów i obiektów. Wreszcie psychologowie dostrzegli w nich idealne wzorce do modelowania i badania procesów informacyjnych w psychice człowieka. Krzemowe (sprzętowe) struktury sztucznych sieci neuronowych rokowały nadzieje na budowę inteligentnych, samouczących się układów scalonych. Takiego uniwersum nauka i praktyka badawcza jeszcze nie doświadczyła.
- W początkowych okresach szczególnie rozległe nadzieje docelowych aplikacji rysowały się przed „elektronowymi mózgiami sieciowymi” w sferze zastosowań komercyjnych. Sieci neuronowe miały rozwiązywać nie tylko skomplikowane problemy naukowe i teoretyczne, ale przede wszystkim codzienne „zadania optymalizacyjne” związane z decyzjami biznesowymi, transportowymi, korporacyjnymi, dotyczące różnych aspektów bezpieczeństwa, kwestii zdrowotnych i wielu innych spraw uwalniając w znacznym stopniu człowieka od ryzyka podejmowanych