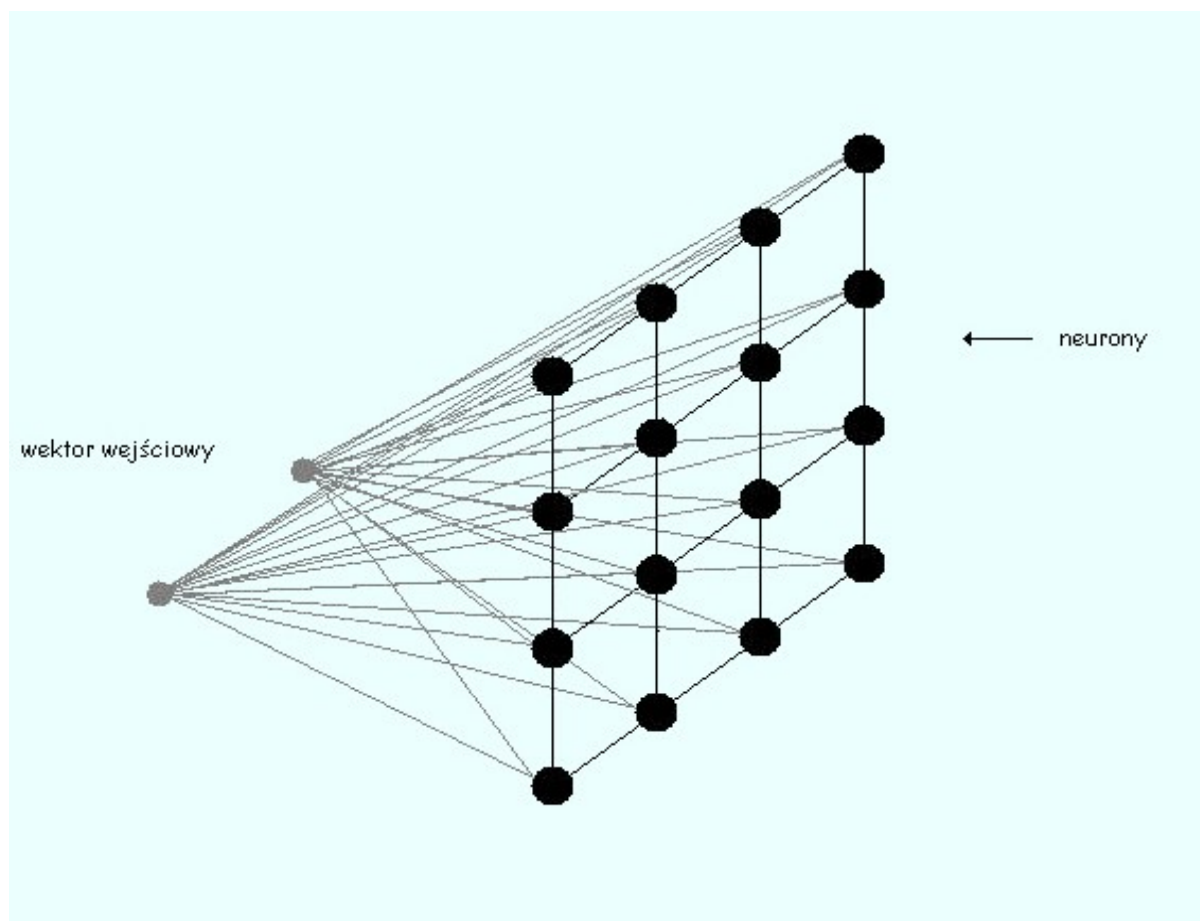


::Czym sa sieci Kohonena?::

Sieci Kohonena są jednym z podstawowych typów sieci samoorganizujących się. Właśnie dzięki zdolności samoorganizacji otwierają się zupełnie nowe możliwości - adaptacja do wcześniej nieznanych danych wejściowych, o których bardzo niewiele wiadomo. Wydaje się to naturalnym sposobem uczenia, który jest używany chociażby w naszych mózgach, którym nikt nie definiuje żadnych wzorców, tylko muszą się one krystalizować w trakcie procesu uczenia, połączonego z normalnym funkcjonowaniem. Sieci Kohonena stanowią synonim całej grupy sieci, w których uczenie odbywa się metodą samoorganizującą typu konkurencyjnego. Polega ona na podawaniu na wejścia sieci sygnałów, a następnie wybraniu w drodze konkurencji zwycięskiego neuronu, który najlepiej odpowiada wektorowi wejściowemu. Dokładny schemat konkurencji i późniejszej modyfikacji wag synaptycznych może mieć różną postać. Wyróżnia się wiele podtypów sieci opartych na konkurencji, które różnią się dokładnym algorytmem samoorganizacji.

::Architektura sieci saorganizującej się::

Bardzo istotną kwestią jest struktura sieci neuronowej. Pojedynczy neuron jest mechanizmem bardzo prostym i przez to niewiele potrafiącym. Dopiero połączenie wielu neuronów ze sobą umożliwia prowadzenie dowolnie skomplikowanych operacji. Ze względu na raczej niewielką wiedzę o faktycznych zasadach funkcjonowania ludzkiego mózgu, powstało wiele różnych architektur, które starają się naśladować budowę i zachowanie poszczególnych fragmentów układu nerwowego. Najczęściej stosuje się w tego typu sieciach architekturę jednokierunkową jednowarstwową. Jest to podyktowane faktem, że wszystkie neurony muszą uczestniczyć w konkurencji na równych prawach. Dlatego każdy z nich musi mieć tyle wejść ile jest wejść całego systemu.



Użyta sieć była siecią o schemacie kwadratowym jak na rysunku o zmiennej liczbie neuronów od 5 do 20.

::Etapy działania::

Funkcjonowanie samoorganizujących się sieci neuronowych odbywa się w trzech etapach:

- konstrukcja
- uczenie
- rozpoznawanie

System, który miałby realizować funkcjonowanie sieci samoorganizującej powinien składać się z kilku podstawowych elementów. Pierwszym z nich jest macierz neuronów pobudzanych przez sygnały wejściowe. Sygnały te powinny opisywać pewne charakterystyczne cechy zjawisk zachodzących w otoczeniu, tak, aby na ich podstawie sieć była w stanie je pogrupować. Informacja o zdarzeniach jest przekładana na bodźce pobudzające neurony. Zbiór sygnałów przekazywanych do każdego neuronu nie musi być identyczny, nawet ich ilość może być różna. Muszą one jednak spełniać pewien warunek, a mianowicie jednoznacznie określać dane zdarzenia.

Kolejną częścią składową sieci jest mechanizm, który dla każdego neuronu określa stopień podobieństwa jego wag do danego sygnału wejściowego oraz wyznacza jednostkę z największym dopasowaniem - zwycięzcę. Obliczenia zaczynamy dla wag równych małym liczbom losowym, przy czym ważne jest, aby nie zachodziła żadna symetria. W trakcie uczenia wagi te są modyfikowane w taki sposób, aby najlepiej odzwierciedlać wewnętrzną strukturę danych wejściowych. Istnieje jednak niebezpieczeństwo, że zwiążą się one z pewnymi wartościami zanim jeszcze grupy zostaną prawidłowo rozpoznane i wtedy trzeba ponawiać uczenie z innymi wagami.

Wreszcie konieczne do przeprowadzenia samoorganizacji jest, aby sieć była wyposażona w zdolność do adaptacji wartości wag neuronu zwycięzcy i jego sąsiadów w zależności od siły, z jaką odpowiedział on na dane wejście. Topologię sieci można w łatwy sposób określić poprzez zdefiniowanie sąsiadów dla każdego neuronu. Załóżmy, że jednostkę, której odpowiedź na dane pobudzenie jest maksymalna, będziemy nazywali "obrazem" tego pobudzenia. Wtedy możemy przyjąć, że sieć jest uporządkowana, jeśli topologiczne relacje między sygnałami wejściowymi i ich obrazami są takie same.

::Algorytm uczenia::

Algorytmem, od którego nazwę wzięła cała klasa sieci są samoorganizujące się mapy Kohonena. Zostały one opisane przez ich twórcę w publikacji "The Self Organising Map". Kohonen zaproponował dwa rodzaje sąsiedztwa: prostokątne i gaussowskie.

$$G(i, x) = \begin{cases} 1 & \text{dla } d(i, w) \leq \lambda \\ 0 & \text{dla } d(i, w) > \lambda \end{cases}$$

$$G(i, x) = \exp\left(-\frac{d^2(i, w)}{2\lambda^2}\right)$$

λ jest promieniem sąsiedztwa, malejącym w czasie.

W użytym algorytmie zastosowano sąsiedztwo gaussowskie. Promień przyjmował wartości: [1.0, 2.0, 3.0]. Natomiast współczynnik uczenia zmieniał się w czasie zgodnie ze wzorem:

$$\eta = \eta_0 * e^{(t/\lambda)}$$

gdzie:

t – iteracja

λ – stała równa liczbie epok pomnożonej przez liczbę przypadków w jednej epoce

Współczynnik uczenia: η_0 zmieniał się w granicach [0.5, 0.01]

Sieć była uczona przez 100 epok

Źródło: http://galaxy.agh.edu.pl/~vlsi/AI/koho_t/

::Wyniki::

Wyniki są przedstawione w innym pliku z racji na bardziej przejrzyste wyświetlanie.

::Analiza::

Ogólnie można powiedzieć, że sieci o mniejszym rozmiarze lepiej spisują się jeśli chodzi o klasyfikację danych zaszumionych, niż sieci o większym rozmiarze. Sieci o mniejszym rozmiarze mają lepsze wyniki jeśli chodzi o powtórność „trafności” w wytypowaniu do jakiej kategorii zaklasyfikować zaszumione dane.

Żadna z sieci nie potrafiła dobrze zaklasyfikować wszystkich zaszumionych przypadków. Maksymalna uzyskana dokładność to 75% dla sieci o rozmiarze 5

```
DATA:
L . . . . . D . . . . . J . . .
E . . . . . B . . . . . C . . .
F . . . . . PR . . . . . G . . .
H . . . . . A . . . . . Q . . .
N . . . . . M . . . . . K . . .
IT . . . . .

TEST DATA:
L . . . . . D . . . . . J . . .
. . . . . B . . . . . . . . .
EF . . . . . PR . . . . . A . . .
CG . . . . . S . . . . .
H . . . . . . . . . . . . . .
OQ . . . . .
N . . . . . M . . . . . K . . .
IT . . . . .

size:5, radius:1.0, learningRate:0.5, endLearnRate:0.1840317, accuracy:75.0
```

Zazwyczaj w danych można wyróżnić pewne grupy liter, które przez wszystkie sieci są uznawane za podobne, są to:

- HK
- MN
- PFER
- DOQ
- CG
- IT

Które wykazują rzeczywiście pewne podobieństwo

Sieci o mniejszym współczynniku uczenia i małym promieniu miały tendencję do skupiania danych wokół siebie

Sieci o większym współczynniku uczenia, bądź te o większym promieniu zazwyczaj pokrywały wynikami większość sieci

.	.	FP
.	E	R	K
.	L	AQ	.	HMN
.	CG	DJO
.	.	BS	.	T
.
.
.
.
.
.
.
.
.
.
.	I

.	.	EFP
.	B	R	HK
.	L	AQQ	M	N
.	.	G	D
.	C	JS	.	T
.
.
.
.
.
.
.
.
.
.
.	I

C	.	.	O	.	.	.	J	.	.	.	G	.	.	.
.	.	.	D	S
.
A	Q	I	.	.	.
.
.	T	.	.	.
L	.	.	.	B
E	H	K
.
F	.	.	P	.	.	R	N	.	.	M

	J	.	.	G
.	C	.	.
.	.	.	.	D
A	Q	S
.	Q	.	.	.	I
.
.	T	.	.
L
.	.	.	B	H	K
.	.	.	.	P
.
EF	R	.	.	.	N	.	.	M

::Wnioski::

- początkowy współczynnik uczenia ma istotny wpływ na „rozrzut” danych, im mniejszy, tym mniejszy „rozrzut”
- początkowa wartość promienia ma istotny wpływ na „rozrzut” danych, im mniejszy tym mniejszy rozrzut
- im mniejszy rozmiar sieci, tym lepiej sieć klasyfikuje zaszumione dane
- każda z sieci wykazuje tendencję do „zbliżania do siebie” danych, które wykazują podobieństwo między sobą