

Krzysztof Kaflowski

IS

Podstawy Sztucznej Inteligencji

„Rozpoznawanie charakterystycznych cech dużych liter alfabetu”

1. Celem ćwiczenia było stworzenie sieci Kohonena wykorzystującej algorytm WTM, który miałby rozpoznawać cechy charakterystyczne dużych liter i przydzielać je do odpowiednich grup.
2. Teoria

Algorytm WTM – Winner Takes Most – jest zbliżony do algorytmu WTA, lecz w tym przypadku poza zwycięskim neuronem również jego sąsiedztwo zmienia wagi. Wraz z kolejnymi iteracjami programu początkowy promień sąsiedztwa oraz początkowy współczynnik uczenia jest zmniejszany.

SOM – Self Organizing Map – Samo organizująca się mapa jest typem sztucznej sieci neuronowej, która jest przeznaczona do nauki bez nauczyciela. Jej celem jest przedstawienie wielowymiarowych danych za pomocą małowymiarowego wyniku, zwykle dwuwymiarowego. W wyniku działania programu powstaje mapa, która grupuje podobne do siebie dane. Znajduje zastosowanie zwłaszcza w przypadkach, gdzie jest dużo danych dla jednego rekordu.

Sieć, która uczy się bez nauczyciela musi sama rozpoznać czego się uczy. Nie jesteśmy w stanie rozpoznać w ten sposób konkretnych liter, lecz jesteśmy w stanie rozpoznawać charakterystyczne cechy liter przez co można je pogrupować na podobne do siebie.

Sieć składa się z jednej warstwy wejściowej oraz jednej warstwy wyjściowej. Warstwa wejściowa przekazuje dane uczące do każdego neuronu warstwy wyjściowej. Warstwa wyjściowa jest to mapa pogrupowanych danych.

W sieci Kohonena neurony nie są ze sobą powiązane.

Analizowane dane muszą być odwzorowane na mapie (rozmiary mapy są zadawane przez użytkownika). Zakładając, że na mapie znajduje się M neuronów. Liczba M jest deklarowana, lub też przyjmuje się domyślnie:

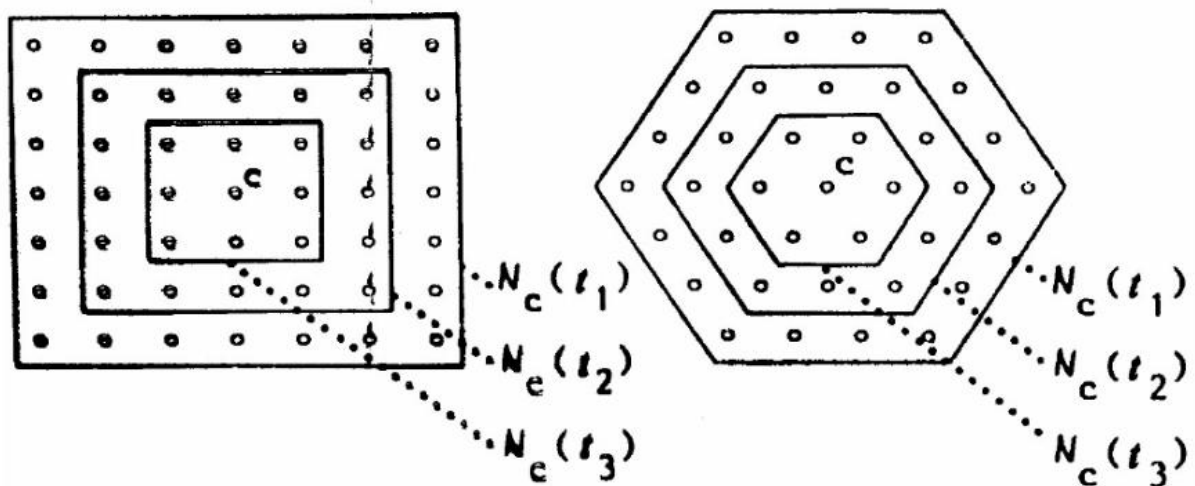
$$M = 5 * \sqrt{N}$$

Gdzie:

- N oznacza liczbę wektorów danych.

Zatem dla naszego przypadku będzie to $5 * \sqrt{35} = 29,5$ co daje w zaokrągleniu 30 neuronów.

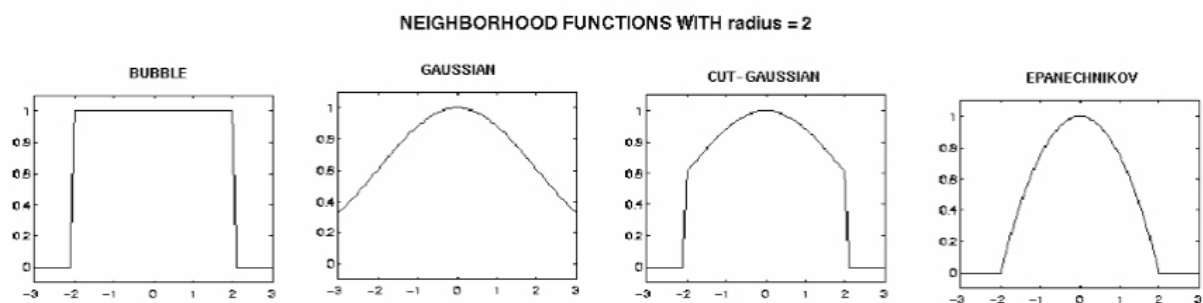
Sąsiedztwo odgrywa bardzo ważną rolę w sieci. Sąsiedztwo to wyznacza się według położenia wektorów referencyjnych na mapie.



Sąsiedztwo w sieciach SOM wyznacza zasięg wpływu zwycięskiego neuronu pobliskie neurony. Sąsiedztwo może być sztywne (0 - nie należy, 1 - należy) lub też wyznaczone funkcją.

Do najpopularniejszych funkcji wyznaczających sąsiedztwo należą funkcje: Bubble, Gaussa, Obcięta Gaussa oraz Epanechnikov

W realizowanym ćwiczeniu zaimplementowano funkcję Gaussa, na mapie w przestrzeni 1D



Nauka sieci

Podczas uczenia się danego rekordu, będzie wyłaniany jeden zwycięski neuron i następnie dostosowywane wagi w nim oraz jego sąsiedztwie. Po skończonym procesie nauki taki neuron będzie rozpoznawał daną klasę obiektów, ponieważ tylko on będzie się aktywował. Zwycięski neuron typowany jest za pomocą miary Euklidesowej. Zwycięża neuron zwracający najmniejszą wartość d dla danej iteracji algorytmu.

$$d(x, w_i) = \|x - w_i\| = \sqrt{\sum_{j=1}^N (x - w_{i,j})^2}$$

X – wektor rekordu podawanego w warstwie wejściowej

W – wektor wag i -tego neuronu

Współczynnik uczenia

Współczynnik uczenia jest obliczany przy każdej iteracji programu. Jest on malejący.

Został wykorzystany poniższy wzór:

$$L(t) = L_0 * e^{(-\frac{t}{\lambda})}$$

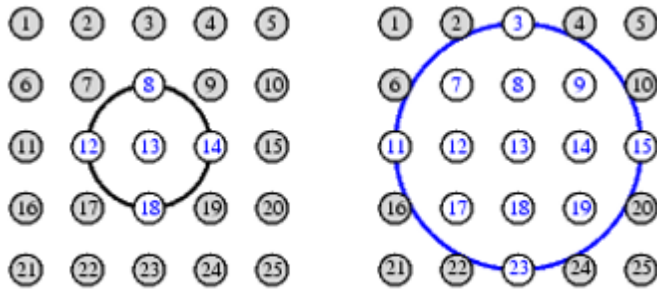
L_0 - Jest początkową wartością współczynnika uczenia

t - Jest numerem epoki uczenia

λ - Stanowi końcową ilość epok uczenia(limit)

Funkcja rozpadu wykładniczego – promień sąsiedztwa

W algorytmie Kohonena, obszar sąsiedztwa kurczy się w kolejnych iteracjach. Funkcja ta może być realizowana poprzez zastosowanie funkcji rozpadu wykładniczego.



Wykorzystano wzór:

$$\sigma(t) = \sigma_0 * e^{(-\frac{t}{\lambda})}$$

σ_0 – Jest początkową średnicą sąsiedztwa (często ustalane jako szerokość całej mapy)

t – Jest numerem epoki uczenia

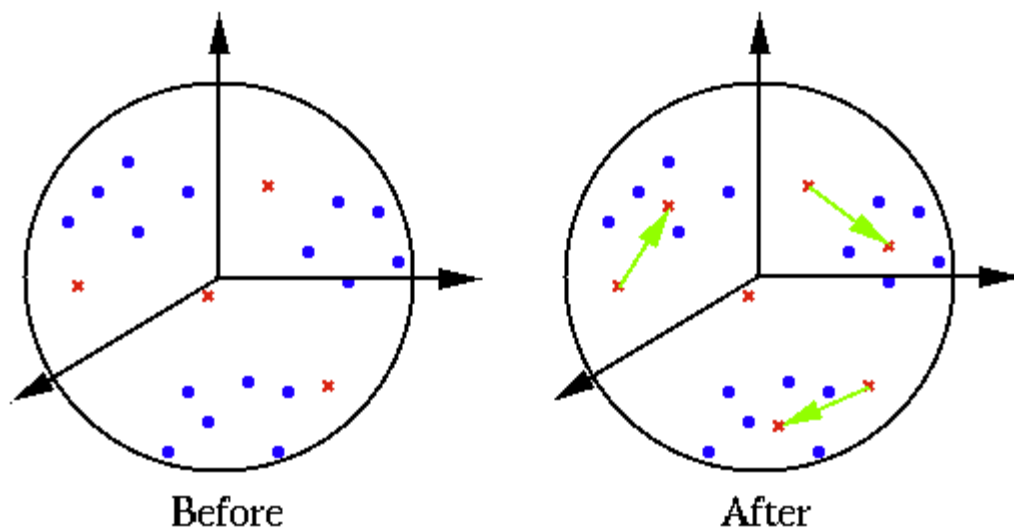
λ – Stanowi końcową ilość epok uczenia(limit)

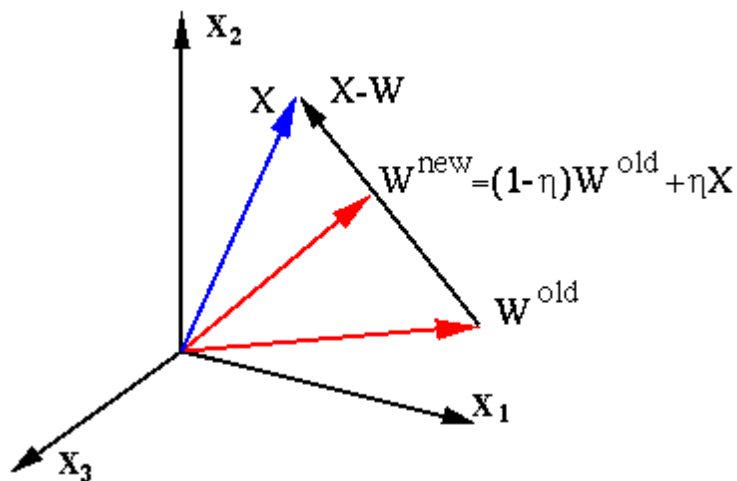
Funkcja theta

Wpływ neuronu zwycięskiego na neurony sąsiadujące

$$\Theta = e^{\frac{\text{dist}^2}{2 * (\sigma(t))^2}}$$

Dist - Jest odległością neuronu dla którego obliczamy zmianę wag, od neuronu zwycięskiego





Reguła uczenia konkurencyjnego powoduje stopniową zmianę kierunku wektora wag poszczególnych jednostek wyjściowych sieci w stronę statystycznie najczęściej występujących wejść.

Regułę uczenia konkurencyjnego można zapisać w postaci uwzględniającej kolejne iteracje uczenia $k, k+1, k+2, \dots$

Normalizacja wektorów

Wykazano że proces samoorganizacji prowadzi zawsze do spójnego podziału przestrzeni cech, gdy choć jeden z wektorów X lub W podlega normalizacji

Jeżeli wektory X są znormalizowane to wektory wag stają się również automatycznie znormalizowane. Badania eksperymentalne potwierdziły potrzebę normalizacji wektorów dla małych wymiarów $N = 2, 3, \dots$. Dla $N > 200$ normalizacja nie odgrywa większej roli.

Modyfikacja wag:

W każdej iteracji wyłaniany jest neuron zwycięski oraz jego sąsiedztwo. Modyfikacja wag następuję zgodnie z podanym wzorem:

$$W(t+1) = W(t) + \Theta(t) * L(t) * (X(t) - W(t))$$

Weryfikacja map cech

Końcowa postać mapy cech obciążona jest dużym wpływem składnika losowego, dlatego po nauczaniu się sieci warto poddać ocenę jakość uzyskanego przez nią odwzorowania. Można tego dokonać, obliczając dla każdego wektora uczącego tzw. błąd kwantyzacji. Jest to odległość pomiędzy wektorem uczącym a wektorem wag neuronu zwycięskiego (najbardziej do niego podobnego):

$$QE_i = d(x_i, w_{i, \text{winner}}) = \min d(x_i, w_i)$$

X – Wektor uczący

W_i, winner – Wektor wag zwycięskiego neuronu

Znając wartości błędów dla wszystkich wektorów uczących, można obliczyć przeciętny błąd kwantyzacji:

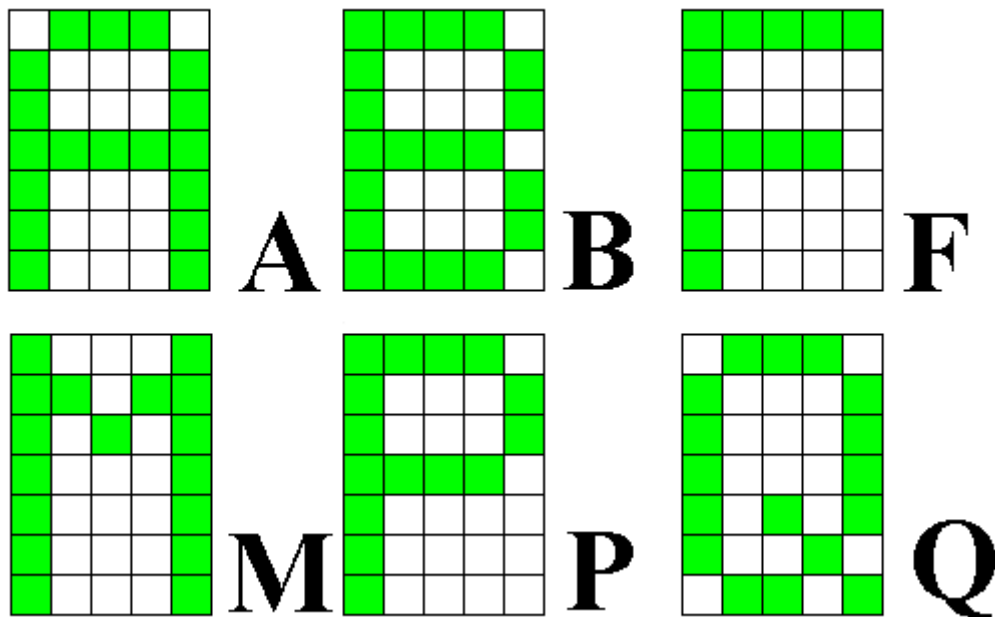
$$AQE = \frac{1}{Z} \sum_{u=1}^Z QE_u$$

Struktura sieci

Stworzona sieć posiada 35 wejść oraz 30 neuronów w warstwie wyjściowej.

3. Dane uczące oraz testujące

Dane uczące oraz testujące są takie same i są zawarte w pliku „learnLetterOnlyBinary.txt”



Siatki liter zostały pobrane ze strony:

<http://www.ai.c-labtech.net/sn/litery.html>

Powyżej przykład wyglądu takich siatek. Kwadraty zakolorowane na zielone są reprezentowane jako wartość 1, a pozostałe jako wartość (-1). Nie zdecydowano się na postać zero-jedynkową ze względu na uwydatnienie charakterystycznych cech w ten sposób.

Modyfikacja wag zachodzi tylko dla zwycięskiego neuronu i jego sąsiedztwa.

Pierwsza linia oznacza liczbę rekordów

Druga linia oznacza liczbę danych

Każdy rekord ma na końcu literę, którą reprezentuje.

Plik „learnLetterOnlyBinary.txt”:

```
26
35
-1 1 1 1 -1 1 -1 -1 -1 1 1 -1 -1 -1 1 1 1 1 1 -1 -1 -1 1 1 -1 -1 -1 1 1
-1 -1 -1 1 A
1 1 1 1 -1 1 -1 -1 -1 1 1 -1 -1 -1 1 1 1 1 1 -1 1 -1 -1 -1 1 1 -1 -1 -1 1 1
1 1 1 -1 B
-1 1 1 1 -1 1 -1 -1 -1 1 1 -1 -1 -1 -1 1 -1 -1 -1 -1 1 -1 -1 -1 -1 1 -1 -1
-1 1 -1 1 1 1 -1 C
1 1 1 1 -1 1 -1 -1 -1 1 1 -1 -1 -1 1 1 -1 -1 -1 1 1 -1 -1 -1 1 1 -1 -1 -1 1
1 1 1 1 -1 D
1 1 1 1 1 1 -1 -1 -1 -1 1 -1 -1 -1 -1 1 1 1 1 -1 1 -1 -1 -1 1 -1 -1 -1 -
1 1 1 1 1 1 E
1 1 1 1 1 1 -1 -1 -1 -1 1 -1 -1 -1 -1 1 1 1 1 -1 1 -1 -1 -1 1 -1 -1 -1 -
1 1 -1 -1 -1 -1 F
-1 1 1 1 -1 1 -1 -1 -1 1 1 -1 -1 -1 -1 1 -1 1 1 1 1 -1 -1 -1 1 1 -1 -1 -1 1
-1 1 1 1 -1 G
1 -1 -1 -1 1 1 -1 -1 -1 1 1 -1 -1 -1 1 1 1 1 1 1 -1 -1 -1 1 1 -1 -1 -1 1
1 -1 -1 -1 1 H
-1 1 1 1 -1 -1 -1 1 -1 -1 -1 1 -1 -1 -1 1 -1 -1 -1 1 -1 -1 -1 -1 -1 1
-1 -1 -1 1 1 1 -1 I
1 1 1 1 1 -1 -1 -1 1 -1 -1 -1 -1 1 -1 -1 -1 -1 1 -1 -1 -1 1 1 -1 -1 -
1 1 -1 1 1 1 -1 J
1 -1 -1 -1 1 1 -1 -1 1 -1 1 -1 1 -1 -1 1 1 -1 -1 1 -1 1 -1 -1 1 -1 -1 1
-1 1 -1 -1 0 1 K
```

```

1 -1 -1 -1 -1 1 -1 -1 -1 -1 1 -1 -1 -1 -1 1 -1 -1 -1 -1 1 -1 -1 -1 -1
-1 -1 -1 1 1 1 1 1 L
1 -1 -1 -1 1 1 1 -1 1 1 1 -1 1 -1 1 1 -1 -1 -1 1 1 -1 -1 -1 1
1 -1 -1 -1 1 M
1 -1 -1 -1 1 1 -1 -1 -1 1 1 1 -1 -1 1 1 -1 1 1 1 -1 -1 -1 1
1 -1 -1 -1 1 N
-1 1 1 1 -1 1 -1 -1 -1 1 1 -1 -1 -1 1 1 -1 -1 -1 1 1 -1 -1 -1
1 -1 1 1 1 -1 O
1 1 1 1 -1 1 -1 -1 -1 1 1 -1 -1 -1 1 1 1 1 -1 1 -1 -1 -1 -1
1 -1 -1 -1 -1 P
-1 1 1 1 -1 1 -1 -1 -1 1 1 -1 -1 -1 1 1 -1 -1 -1 1 1 -1 -1 1 -1
-1 1 1 -1 1 Q
1 1 1 1 -1 1 -1 -1 -1 1 1 -1 -1 -1 1 1 1 1 -1 1 -1 1 -1 -1 1 -1
-1 -1 -1 1 R
-1 1 1 1 -1 1 -1 -1 -1 1 1 -1 -1 -1 -1 1 1 1 -1 -1 -1 -1 1 1 -1 -1
1 -1 1 1 1 -1 S
1 1 1 1 1 -1 -1 1 -1 -1 -1 -1 1 -1 -1 -1 -1 1 -1 -1 -1 -1 1 -
1 -1 -1 -1 1 -1 -1 T
1 -1 -1 -1 1 1 -1 -1 -1 1 1 -1 -1 -1 1 1 -1 -1 -1 1 1 -1 -1 -1
1 -1 1 1 1 -1 U
1 -1 -1 -1 1 1 -1 -1 -1 1 1 -1 -1 -1 1 1 -1 -1 -1 1 -1 1 -1 1
-1 -1 -1 1 -1 -1 V
1 -1 -1 -1 1 1 -1 -1 -1 1 1 -1 -1 -1 1 1 -1 -1 -1 1 1 -1 1 -1 1
-1 1 -1 1 -1 W
1 -1 -1 -1 1 1 -1 -1 -1 1 -1 1 -1 -1 -1 1 -1 -1 1 -1 1 -1 -1 -
1 1 1 -1 -1 1 X
1 -1 -1 -1 1 1 -1 -1 -1 1 -1 1 -1 -1 -1 1 -1 -1 -1 1 -1 -1 -1
-1 -1 -1 -1 1 -1 -1 Y
1 1 1 1 1 -1 -1 -1 -1 1 -1 -1 -1 1 -1 -1 -1 1 -1 -1 -1 1 -1 -
1 -1 1 1 1 1 1 Z

```

4. Wyniki działania programu

W ramach przetestowania programu wykonano następujące typy testów:

- Zmienny początkowy współczynnik uczenia
- Zmienny początkowy promień sąsiedztwa
- Zmienna ilość epok

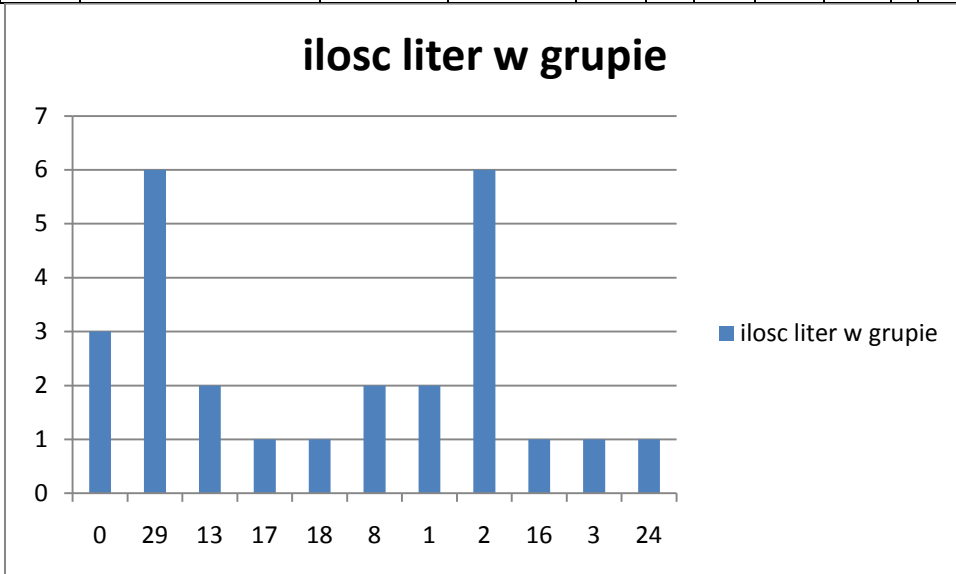
Zmienny początkowy współczynnik uczenia:

początkowy współczynnik uczenia	początkowy promień sąsiedztwa	ilość epok
0,01	40	10000

litera	błąd kwantyzacji
A	1.1189638435463856
B	1.179318140276365
C	1.155545287800954
D	1.1965594696780333
E	1.134104526487195
F	1.1144488580036844
G	1.1438988356522528
H	1.1020294950654899
I	0.9409876249995857
J	1.081571200170631
K	0.9916461026140552

L	1.1209299933875523
M	1.021590570920771
N	1.0583508348066504
O	1.1526557904989485
P	1.145047746997978
Q	1.093318505218717
R	1.0788861506331322
S	1.0915501830232404
T	0.9533709467855376
U	1.1601009087488994
V	1.04041705197275
W	1.0866326991733075
X	0.9656761791469273
Y	0.9669939281575106
Z	1.0205730422224282

	PODSUMOWANIE:									
ID:	0	LLTERY:	A,	F,	R,					3
ID:	29	LLTERY:	B,	D,	J,	O,	Q,	S,		6
ID:	13	LLTERY:	C,	G,						2
ID:	17	LLTERY:	E,							1
ID:	18	LLTERY:	H,							1
ID:	8	LLTERY:	I,	T,						2
ID:	1	LLTERY:	K,	L,						2
ID:	2	LLTERY:	M,	N,	U,	V,	W,	X,		6
ID:	16	LLTERY:	P,							1
ID:	3	LLTERY:	Y,							1
ID:	24	LLTERY:	Z,							1



Błąd globalny: 1.0812964237073444

początkowy współczynnik uczenia	początkowy	ilość epok
---------------------------------	------------	------------

	promien sasiedztw a	
0,1	40	10000

litera

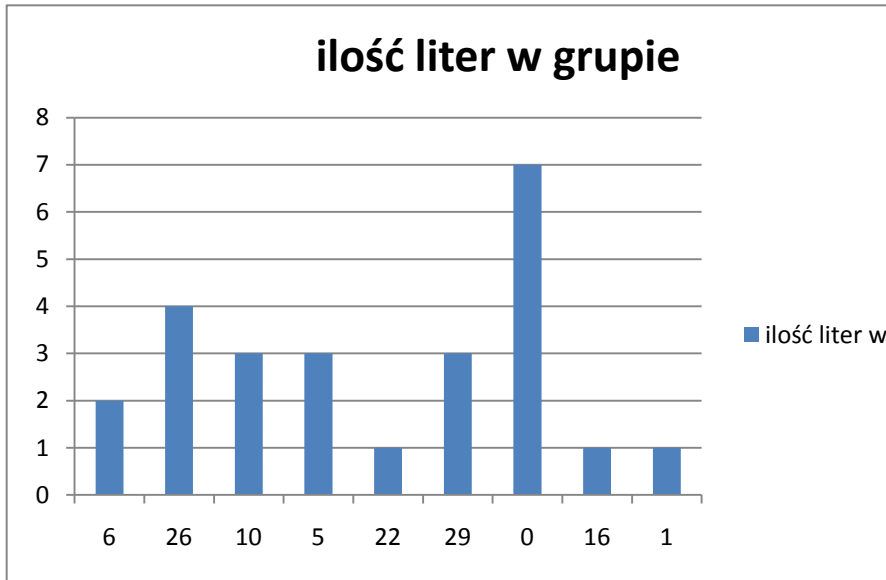
błąd kwantyzacji

	1.0932116507649
A	108
	1.1432291208778
B	562
	1.1427341433023
C	421
	1.1776558924069
D	565
	1.1082563524665
E	217
	1.1035903230303
F	2
	1.1268405973243
G	08
	1.0998442543758
H	93
	0.9471789560223
I	89
	1.0792780482305
J	733
	0.9936693703626
K	763
	1.1172821627309
L	024
	1.0252498260817
M	75
	1.0623377467233
N	93
	1.1436396166318
O	166
	1.1253596048066
P	876
	1.0818455379151
Q	204
	1.0637525239964
R	416
	1.0785765119137
S	67
	0.9601207421458
T	035
	1.1636155146114
U	648
	1.0477249016875
V	376
	1.0871648607422
W	62

X
Y
Z

0.9708532658446
392
0.9829045611248
571
1.0287406935196
712

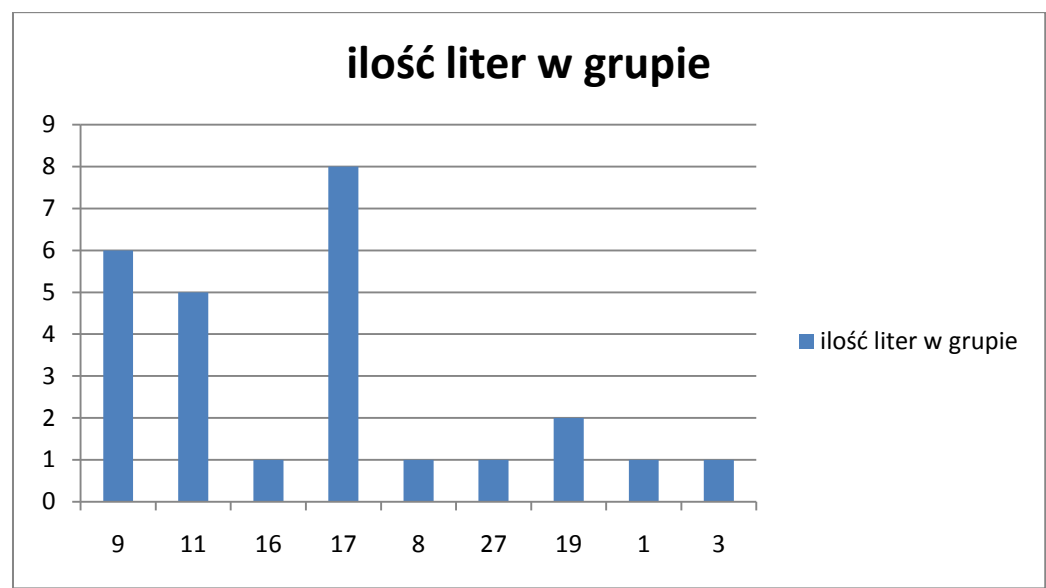
	PODSUMOWANIE										
	:										
ID		LLTERY									
:	6	:	A,	P,							2
ID		LLTERY									
:	26	:	B,	D,	O,	S,					4
ID		LLTERY									
:	10	:	C,	G,	Q,						3
ID		LLTERY									
:	5	:	E,	F,	R,						3
ID		LLTERY									
:	22	:	H,								1
ID		LLTERY									
:	29	:	I,	J,	T,						3
ID		LLTERY									
:	0	:	K,	L,	M	N,	U	W	X,	Y,	7
ID		LLTERY									
:	16	:	V,								1
ID		LLTERY									
:	1	:	Z,								1



początkowy współczynnik uczenia	początkowy promień sąsiedztwa	ilość epok
0,5	40	10000

litera	błąd kwantyzacji
A	1.01162283761816
B	1.0562949043460674
C	1.0847633280381006
D	1.0945088760303043
E	1.0576191032852
F	1.058464684942648
G	1.0480033191733311
H	1.0626308879306288
I	0.9519880090299367
J	1.0605271111420518
K	0.9847567122578907
L	1.1000625092520786
M	0.9988381706048025
N	1.0613127607793098
O	1.0648227353185589
P	1.0592835818111737
Q	1.0290286000152324
R	1.012462548033946
S	1.0270202068494383
T	0.9748495206860434
U	1.1088288498672334
V	1.0267140792720495
W	1.0556335633109313
X	0.9910238576232273
Y	1.0137511698958064
Z	1.0478708319267942

	PODSUMOWANIE:										
ID:	9	LLTERY:	A,	B,	E,	F,	P,	R,			6
ID:	11	LLTERY:	C,	G,	O,	Q,	S,				5
ID:	16	LLTERY:	D,								1
ID:	17	LLTERY:	H,	K,	L,	M,	N,	U,	V,	W,	8
ID:	8	LLTERY:	I,								1
ID:	27	LLTERY:	J,								1
ID:	19	LLTERY:	T,	Z,							2
ID:	1	LLTERY:	X,								1



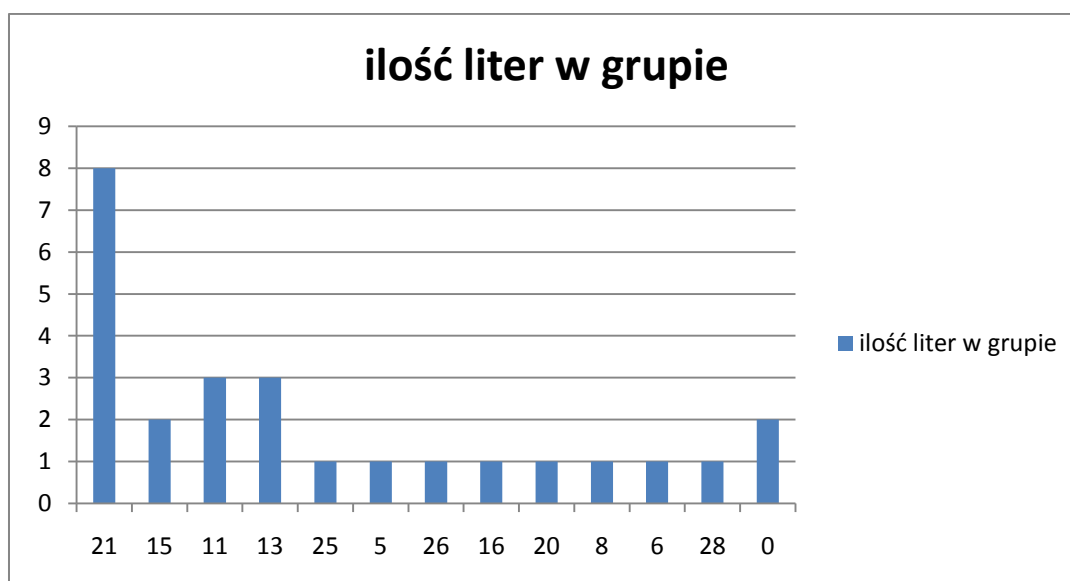
Błąd globalny: 1.0401031830400362

początkowy współczynnik uczenia	początkowy promień sąsiedztwa	ilość epok
0,95	40	10000

litera	błąd kwantyzacji
A	0.9397304209556497
B	0.9915434932218355
C	1.0221983975741145
D	1.009389217236264
E	0.988056760559812
F	1.006266157295379
G	0.992258938104539
H	0.9818053807733828
I	0.8919850322808169
J	1.021749663386206
K	0.945769006500572
L	1.0515436342305189
M	0.9456241568572599
N	0.9889413246756933
O	0.9867768935536052
P	0.9939333705903001
Q	0.9658664324211035
R	0.9636881854082233
S	0.9992296678799187
T	0.9271515792033376
U	1.053861897089809
V	0.9818681611514649

W 1.0124246490232933
X 1.049541571331701
Y 1.0467491042239767
Z 1.015324995051241

PODSUMOWANIE:											
ID:	21	LLTERY:	A,	F,	K,	L,	M,	P,	U,	V,	8
ID:	15	LLTERY:	B,	G,							2
ID:	11	LLTERY:	C,	E,	J,						3
ID:	13	LLTERY:	D,	O,	W,						3
ID:	25	LLTERY:	H,								1
ID:	5	LLTERY:	I,								1
ID:	26	LLTERY:	N,								1
ID:	16	LLTERY:	Q,								1
ID:	20	LLTERY:	R,								1
ID:	8	LLTERY:	S,								1
ID:	6	LLTERY:	T,								1
ID:	28	LLTERY:	X,								1
ID:	0	LLTERY:	Y,	Z,							2



Błąd globalny: 0.9912799265607701

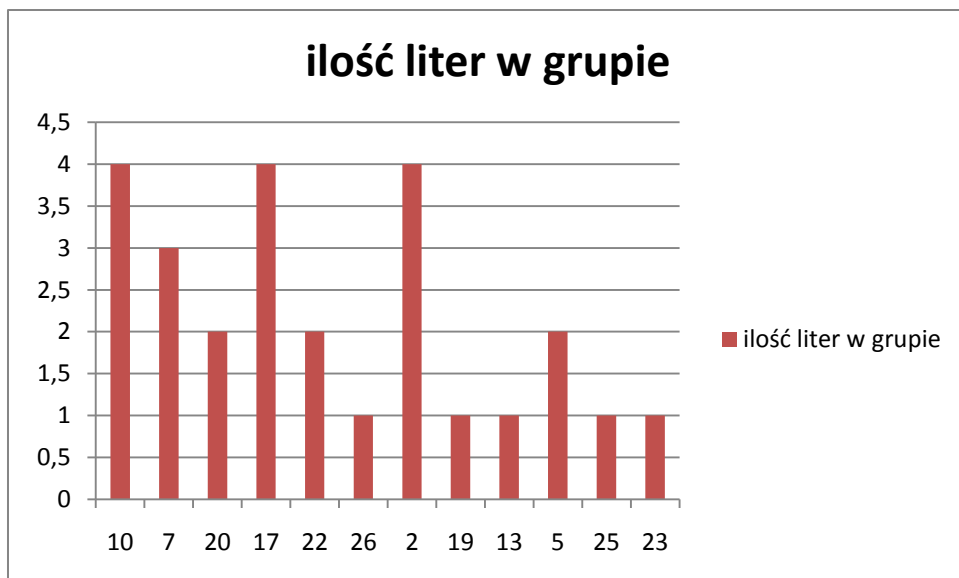
Zmienny początkowy promień sąsiedztwa:

początkowy współczynnik uczenia	początkowy promień sąsiedztwa	ilosc epok
0,1	5	10000

litera	błąd kwantyzacji
A	1.0959207790533647
B	1.148627008422605
C	1.1439747613786073
D	1.170022702054167

E	1.1084436556960342
F	1.094343003882383
G	1.1137469901993788
H	1.1068024810680033
I	0.9324678882032545
J	1.07015821060754
K	0.9805907697226822
L	1.110514312635648
M	1.018960992545984
N	1.0537536460268129
O	1.1222668021052715
P	1.126825825388371
Q	1.0766684953752497
R	1.0687135353168464
S	1.072618033060401
T	0.9463552428439117
U	1.134617594566289
V	1.0356455020121864
W	1.0754615275970278
X	0.9578332242307958
Y	0.9580048266902625
Z	1.0059564524963613

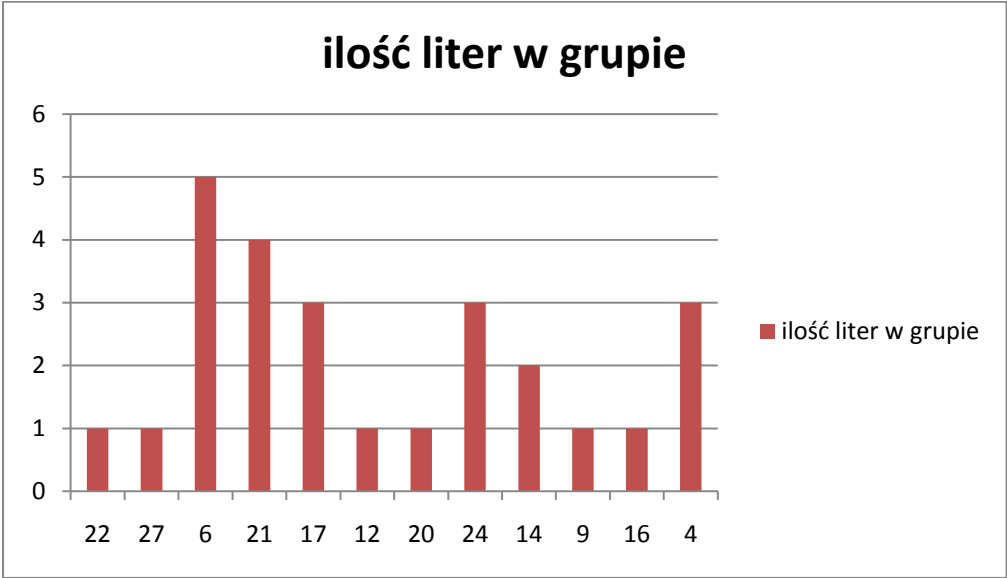
	PODSUMOWANIE:						
ID:	10	LLTERY:	A,	G,	O,	Q,	4
ID:	7	LLTERY:	B,	D,	J,		3
ID:	20	LLTERY:	C,	S,			2
ID:	17	LLTERY:	E,	F,	P,	R,	4
ID:	22	LLTERY:	H,	M,			2
ID:	26	LLTERY:	I,				1
ID:	2	LLTERY:	K,	L,	T,	Y,	4
ID:	19	LLTERY:	N,				1
ID:	13	LLTERY:	U,				1
ID:	5	LLTERY:	V,	X,			2
ID:	25	LLTERY:	W,				1
ID:	23	LLTERY:	Z,				1



początkowy współczynnik uczenia	początkowy promień sąsiedztwa	ilość epok
0,1	10	10000

litera	błąd kwantyzacji
A	1.0878195158806714
B	1.1450963328412411
C	1.13042444420784
D	1.177498357777763
E	1.1038593132701837
F	1.0972868427615505
G	1.122700610330329
H	1.0980440943482876
I	0.9183751478306692
J	1.0701283980405913
K	0.9791876445588074
L	1.1092441015676784
M	1.016304905335247
N	1.0513493715731093
O	1.132998179828489
P	1.127410069080803
Q	1.0885739802139565
R	1.0671641486609253
S	1.0660839955956338
T	0.9384366217212027
U	1.1353529306510444
V	1.0341073596671033
W	1.0747413958798324
X	0.9521570308795159
Y	0.9597519666872537
Z	0.9995485112083465

	PODSUMOWANIE:							
ID:	22	LLTERY:	A,					1
ID:	27	LLTERY:	B,					1
ID:	6	LLTERY:	C,	I,	S,	T,	Z,	5
ID:	21	LLTERY:	D,	G,	J,	O,		4
ID:	17	LLTERY:	E,	P,	R,			3
ID:	12	LLTERY:	F,					1
ID:	20	LLTERY:	H,					1
ID:	24	LLTERY:	K,	X,	Y,			3
ID:	14	LLTERY:	L,	N,				2
ID:	9	LLTERY:	M,					1
ID:	16	LLTERY:	Q,					1
ID:	4	LLTERY:	U,	V,	W,			3



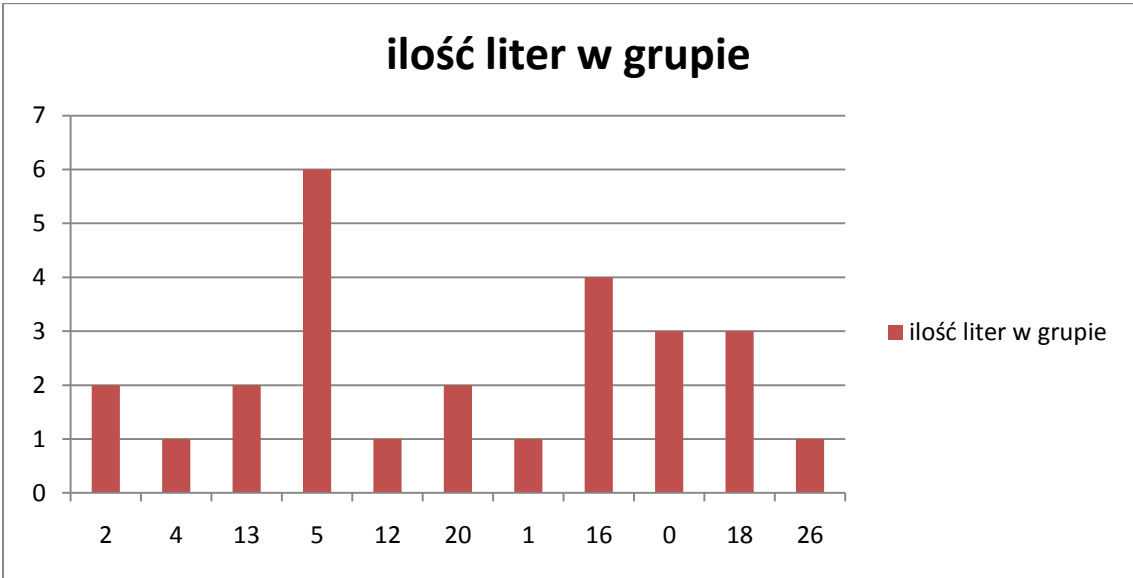
Błąd globalny: 1.064755587323003

początkowy współczynnik uczenia	początkowy promień sąsiedztwa	ilość epok
0,1	20	10000

litera	błąd kwantyzacji
A	1.0830464653306069
B	1.1360945460710132
C	1.141207387323912
D	1.1652333150157477
E	1.1032173838834474
F	1.0914326271884582
G	1.1195226054087213
H	1.0881327075455411
I	0.9367868452875836
J	1.0729249344883747
K	0.9788477614374196

L	1.106184130009721
M	1.0113685193386577
N	1.0490470324680783
O	1.1231695613571178
P	1.1188546751290025
Q	1.0733873212004281
R	1.0641426618114032
S	1.0699934109046068
T	0.9440720202256855
U	1.1466981611666094
V	1.0338444691307302
W	1.0766513509884343
X	0.9552004224262566
Y	0.9580864140565936
Z	1.0107911110368695

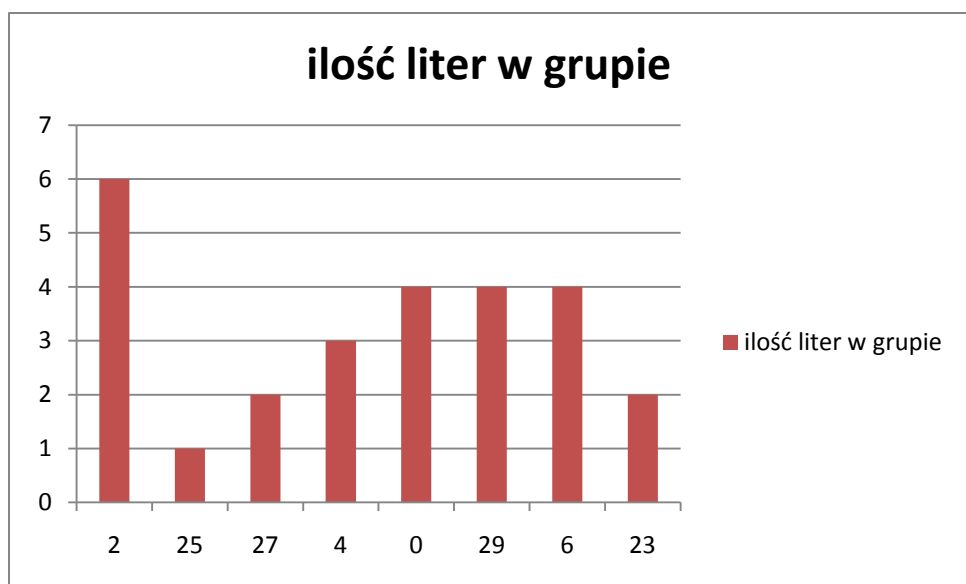
	PODSUMOWANIE:								
ID:	2	LLTERY:	A,	P,					2
ID:	4	LLTERY:	B,						1
ID:	13	LLTERY:	C,	S,					2
ID:	5	LLTERY:	D,	G,	J,	O,	Q,	U,	6
ID:	12	LLTERY:	E,						1
ID:	20	LLTERY:	F,	R,					2
ID:	1	LLTERY:	H,						1
ID:	16	LLTERY:	I,	L,	T,	Z,			4
ID:	0	LLTERY:	K,	X,	Y,				3
ID:	18	LLTERY:	M,	N,	V,				3
ID:	26	LLTERY:	W,						1



początkowy współczynnik uczenia	początkowy promień sąsiedztwa	ilość epok
0,1	60	10000

litera	błąd kwantyzacji
A	1.0853900827579759
B	1.1411720007989157
C	1.1483811967377435
D	1.1695317858391874
E	1.1192842986203682
F	1.1070089821334645
G	1.1181137594974289
H	1.0929177675672015
I	0.9452790543843154
J	1.083582375397781
K	0.9898214032363224
L	1.1126143327358193
M	1.0218853319220262
N	1.0606111036217316
O	1.130083720193014
P	1.132400951649766
Q	1.0711816406653538
R	1.0696820338335102
S	1.0795499587812065
T	0.9609136407547124
U	1.1508638866104473
V	1.0445765519589048
W	1.0826961574092777
X	0.9783913175793298
Y	0.9848512642940337
Z	1.0245245460998933

	PODSUMOWANIE:								
ID:	2	LLTERY:	A,	B,	D,	G,	O,	Q,	6
ID:	25	LLTERY:	C,						1
ID:	27	LLTERY:	E,	F,					2
ID:	4	LLTERY:	H,	P,	R,				3
ID:	0	LLTERY:	I,	J,	S,	Z,			4
ID:	29	LLTERY:	K,	L,	M,	X,			4
ID:	6	LLTERY:	N,	U,	V,	W,			4
ID:	23	LLTERY:	T,	Y,					2



Zmienna ilość epok:

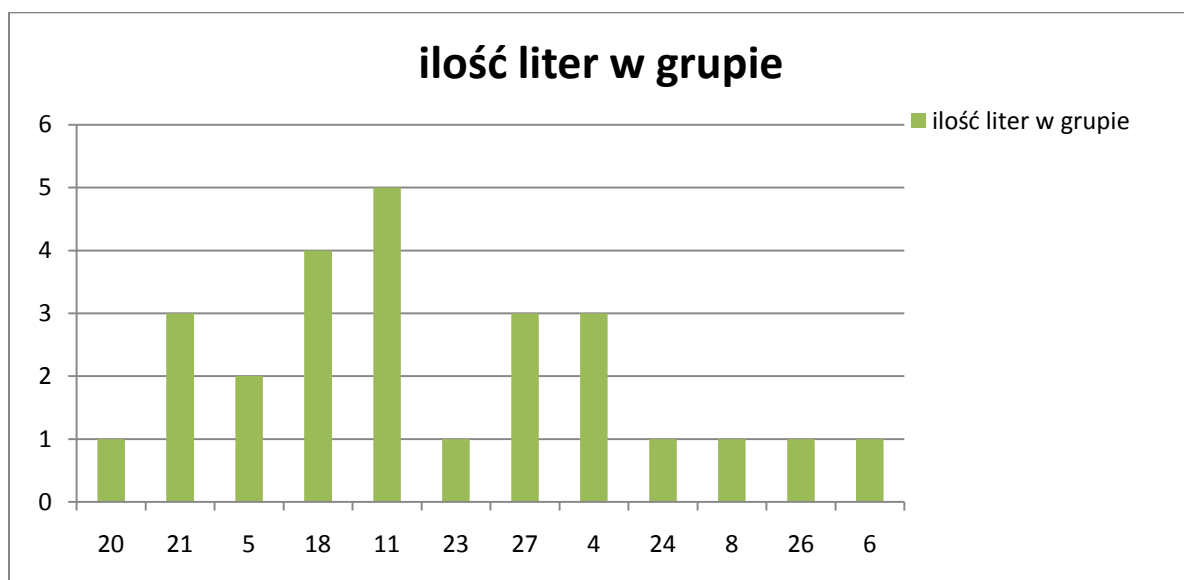
początkowy współczynnik uczenia	początkowy promień sąsiedztwa	ilość epok
0,1	40	100

litera	błąd kwantyzacji
A	1.0967827434689745
B	1.1382459430090726
C	1.1455135094389395
D	1.1758675943074053
E	1.1115110987838168
F	1.1044426513977716
G	1.1259095956790237
H	1.1007902957352484
I	0.9416855117043855
J	1.078120692602882
K	0.9977417570571576
L	1.1207359879756815
M	1.0248852590874709
N	1.0618757190017265
O	1.1422010013088444
P	1.124021882467922
Q	1.0879770873127268
R	1.0733789402167548
S	1.0779918076280963
T	0.9554845184850226
U	1.1569004400674558
V	1.0469286441731909
W	1.08182601277199
X	0.9722785406232602
Y	0.9811517742751416

Z

1.0181211334325018

	PODSUMOWANIE:							
ID:	20	LLTERY:	A,					1
ID:	21	LLTERY:	B,	D,	G,			3
ID:	5	LLTERY:	C,	J,				2
ID:	18	LLTERY:	E,	F,	P,	R,		4
ID:	11	LLTERY:	H,	L,	M,	N,	U,	5
ID:	23	LLTERY:	I,					1
ID:	27	LLTERY:	K,	V,	W,			3
ID:	4	LLTERY:	O,	Q,	S,			3
ID:	24	LLTERY:	T,					1
ID:	8	LLTERY:	X,					1
ID:	26	LLTERY:	Y,					1
ID:	6	LLTERY:	Z,					1



Błąd globalny: 1.0747065439235564

początkowy współczynnik uczenia	początkowy promień sąsiedztwa	ilość epok
0,1	40	1000

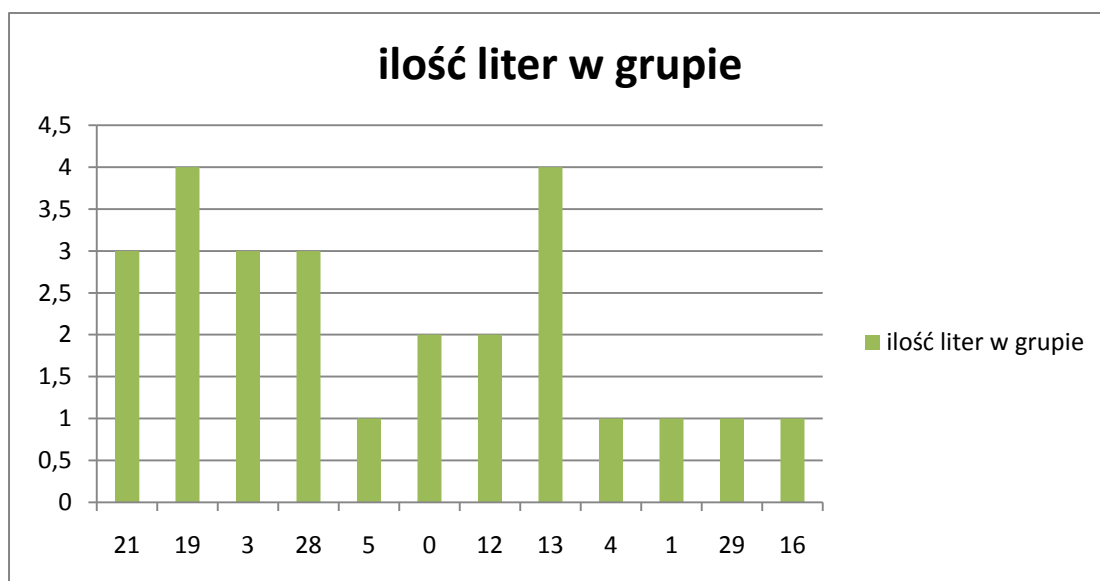
litera

błąd kwantyzacji

A	1.0916824694003504
B	1.1488653735695609
C	1.1462480335908216
D	1.1820358557755886
E	1.1040748407921717
F	1.1039449397018575
G	1.1278574361536382
H	1.095646055383085

I	0.9459069680149694
J	1.0837192068699848
K	0.992853906050872
L	1.1108805062029623
M	1.0218429504500102
N	1.0647238875996023
O	1.1389561744412873
P	1.1249958270380913
Q	1.0829513822563201
R	1.0620506042739406
S	1.0821721918892475
T	0.9622837183474992
U	1.1498762082342302
V	1.0371566458873416
W	1.084926742755139
X	0.9756990504235664
Y	0.9819743464740299
Z	1.0241058906790594

	PODSUMOWANIE:						
ID:	21	LLTERY:	A,	B,	P,		3
ID:	19	LLTERY:	C,	G,	I,	Q,	4
ID:	3	LLTERY:	D,	O,	S,		3
ID:	28	LLTERY:	E,	F,	R,		3
ID:	5	LLTERY:	H,				1
ID:	0	LLTERY:	J,	T,			2
ID:	12	LLTERY:	K,	L,			2
ID:	13	LLTERY:	M,	V,	W,	Y,	4
ID:	4	LLTERY:	N,				1
ID:	1	LLTERY:	U,				1
ID:	29	LLTERY:	X,				1
ID:	16	LLTERY:	Z,				1

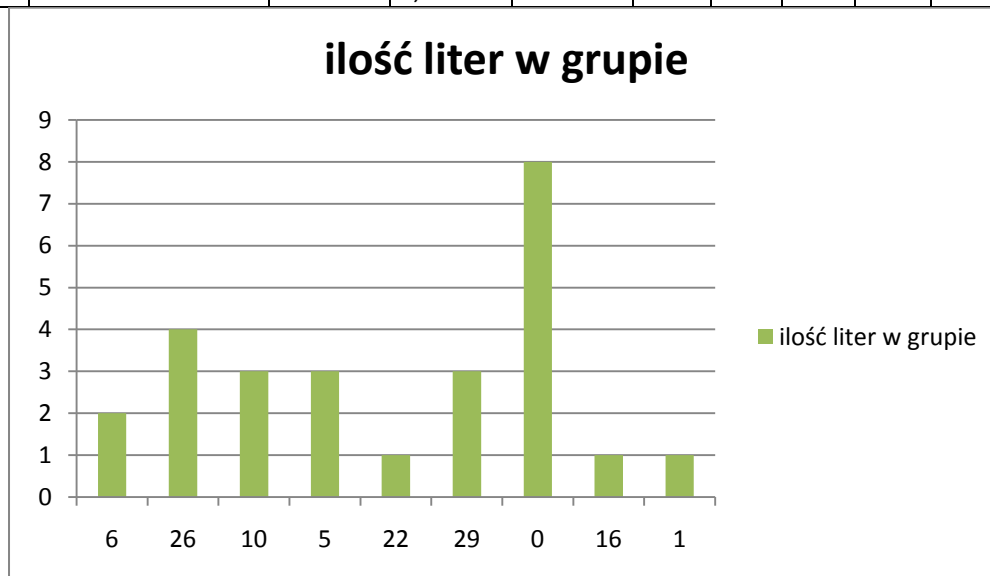


Błąd globalny: 1.0741319697021243

początkowy współczynnik uczenia	początkowy promień sąsiedztwa	ilość epok
0,1	40	10000

litera	błąd kwantyzacji
A	1.0932116507649108
B	1.1432291208778562
C	1.1427341433023421
D	1.1776558924069565
E	1.1082563524665217
F	1.10359032303032
G	1.126840597324308
H	1.099844254375893
I	0.947178956022389
J	1.0792780482305733
K	0.9936693703626763
L	1.1172821627309024
M	1.025249826081775
N	1.062337746723393
O	1.1436396166318166
P	1.1253596048066876
Q	1.0818455379151204
R	1.0637525239964416
S	1.078576511913767
T	0.9601207421458035
U	1.1636155146114648
V	1.0477249016875376
W	1.087164860742262
X	0.9708532658446392
Y	0.9829045611248571

	PODSUMOWANIE:										
ID:	6	LLTERY:	A,	P,							2
ID:	26	LLTERY:	B,	D,	O,	S,					4
ID:	10	LLTERY:	C,	G,	Q,						3
ID:	5	LLTERY:	E,	F,	R,						3
ID:	22	LLTERY:	H,								1
ID:	29	LLTERY:	I,	J,	T,						3
ID:	0	LLTERY:	K,	L,	M,	N,	U,	W,	X,	Y,	8
ID:	16	LLTERY:	V,								1
ID:	1	LLTERY:	Z,								1



Błąd globalny: 1.075179106909265

5. Analiza działania programu

Wszystkie trzy zmienne czynniki mają duży wpływ na działanie programu

Przy zwiększaniu początkowego współczynnika uczenia tworzyło się dużo grup. Grupy te miały jednak po jednym elemencie, a z kolei jedna z grup miała bardzo wiele elementów. Ma to swoje plusy ponieważ można określić dokładnie niektóre litery, ale z drugiej strony jedna z grup ma bardzo dużo wartości niekoniecznie podobnych do siebie.

Przy zwiększaniu początkowego promienia sąsiedztwa tworzyło się coraz mniej grup i w dodatku o równiejszej ilości elementów. Nie ma już dużej przewagi jednej grupy względem innych.

Przy zwiększaniu ilości epok zmniejszała się również liczba grup. Najlepszym tutaj rozwiązaniem nie była skrajnie mała lub skrajnie niska wartość, lecz coś po środku. Dla 10 000 tysięcy epok stworzyła się jedna grupa, która przeważa nad innymi, a dla 1000 epok, grup było więcej, ale były równiejsze co według mnie jest na plus.

Najlepsze ustawienia dla programu

0,1 – początkowy współczynnik uczenia

40 – początkowy promień sąsiedztwa

1000 – ilość epok nauczania

6. Wnioski

- Dzięki Sieci Kohonena z użyciem algorytmu WTM możemy rozpoznawać charakterystyczne cechy danych wejściowych i grupować je względem ich podobieństwa
- Algorytm WTM dobrze sprawdza się przy dużych wymiarach danych wejściowych, które później można reprezentować w małych wymiarach
- Aby sieć poprawnie działała należy dobrać odpowiednie współczynniki początkowe
- Każdy ze współczynników miał duży wpływ na działanie programu

7. Listing kodu programu

Kod programu, dane uczące, dane testujące i to sprawozdanie znajduje się pod linkiem:

<https://github.com/kkaflows/PSI---Scenariusz-6>