

Lekcja 5: Sieci Kohonena i sieci ART.

Problem grupowania danych

S. Hoa Nguyen

1 Materiał

Sieci Kohonena (Sieć samo-organizująca)

Charakterystyka sieci: Jednowarstwowa jednokierunkowa sieć. Funkcja aktywacji: ciągła. Na ogół neurony na warstwie wyjściowej tworzą pewną topologię, np. mapa kwadratowa

Algorytm uczenia

- inicjalizacja wagi sieci,
- pobranie przykładu uczącego,
- obliczenie odległości wektora wejściowego do wag każdego z neuronów,
- wybranie neuronu zwycięzcy (wygrywającego) dla którego odległość wag od wektora wejściowego jest najmniejsza (np. Euklidesowa miara odległości, iloczyn skalarny wektorów),
- zmiana wartości poszczególnych wag tego neuronu:

$$W_{new} = W_{old} + \eta * (X - W_{old})$$

- powtórzenie kroków 2-5 dla wszystkich przykładów uczących.

Sieci ART (Adaptive Resonance Theory)

Charakterystyka sieci: Dwuwarstwowa sieć. Funkcja aktywacji: liniowa. Wektor wejściowy: binarny.

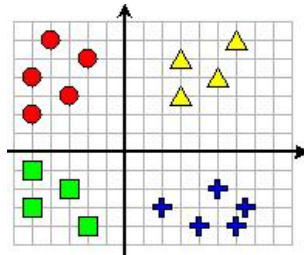
Algorytm uczenia

- inicjalizacja wag sieci: Wektor obrazu docelowego: $V_i = [1]$ dla $i = 1, 2, \dots, k$, wektor wag neuronu $W_i = \frac{1}{1+n}$
- pobranie przykładu uczącego,
- obliczenie odległości wektora wejściowego do wag każdego z neuronów,
- wybranie neuronu zwycięzcy (wygrywającego) - niech nim będzie neuron d

- e) obliczenie współczynnik podobieństwa: $\rho = \frac{\sum v_{di}x_i}{\sum x_i}$
- f) jeśli ρ jest większe niż *współczynnik czułości* to
- g) aktualizacja wag tego neuronu: $V_d^{new} = V_d^{old} \cap X, W_d^{new} = \frac{V_d^{new}}{\alpha + V_d^{new} \diamond X}$
- h) powtórzenie kroków 2-5 dla wszystkich przykładów uczących.

2 Zadania podstawowe

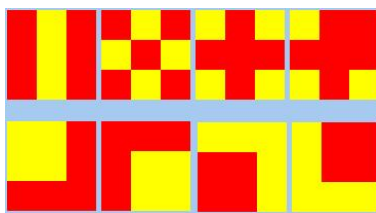
Zadanie 1 Użyto sieci Kohonena do grupowania punktów na płaszczyźnie. Na podstawie przykładów przyporządkowania grupy odpowiednim punktom (patrz rysunek 1), podaj parametry użytej wewnątrz sieci (rozmiar, wagi, funkcje aktywacji itd.).



Rysunek 1: Zbiór punktów do zadania 1.

Zadanie 2 (Grupy drukarek) W celu podziału zbioru drukarek na cztery grupy o podobnych cechach, zastosowano sieć Kohonena (sieć samoorganizująca). W zbiorze *Opisy_Drukarek* są parametry drukarek. Zaprojektować sieć pozwalającą podzielić tego zbioru danych (liczba wejść, liczba neuronów). Wykonać następujące czynności w Excelu:

- a) Symulować działanie sieci dla przykładów treningowych używając strategii WTA (*Winner Takes All*):
 - Unormalizuj wektory cech.
 - Wyznacz dla każdego wektora treningowego neuron o najbliższym wektorze wag.
 - Dla neuronu zwycięskiego aktualizuj wagi.
- b) Narysować położenie danych wejściowych (każda para opisów drukarki odpowiada jednemu punktu w przestrzeni R^2)
- c) Narysować grupy wektorów wejściowych wokół tego samego neuronu.



Rysunek 2: Zbiór punktów do zadania 3.

Zadanie 3 (Sieć ART) Przeprowadź proces uczenia sieci ART dla obrazów podanych na rysunku 2 przyjmując kolejno jako wartość progu czujności wartości 0.4, 0.6 a następnie 0.8. Z ilu neuronów składa się warstwa przetwarzająca każdej z "nauczonych sieci"?

Zadanie 4 (Grupy kwiatów (3 pkt)) Symulować działania następujących algorytmu grupowania danych

- algorytmy 3-średnich i 3-centroidów
- sieć Kohonena dla zbioru danych z opisami Irisów (Data/Iris).

w celu podziału zbioru kwiatków na trzy grupy. Czy otrzymane grupy są zgodne z klasami, do których należą kwiatki?