

Sprawozdanie 5

Celem projektu było poznanie budowy i działanie sieci Kohonena.

Sieć jednokierunkowa to sieć neuronowa, w której nie występuje zjawisko sprzężenia zwrotnego, czyli każdy sygnał przechodzi przez każdy neuron tylko raz.

Sieć Kohonena to podstawowy typ sieci samoorganizującej. Różni się od innych tego typu sieci tym, że nie dostaje ona wzorca, do którego ma dążyć. Dostaje jedynie nieprzetworzone informacje, które sama interpretuje. Uczy się ona w taki sposób, że wybiera na drodze konkurencji neuron, który najlepiej odpowiada wektorowi wejściowemu. Nauka tej sieci przypomina proces nauki u ludzi.

Algorytm uczenia WTA:

Przyjmujemy losowe, znormalizowane wartości wag poszczególnych neuronów. Po podaniu pierwszego wektora wejściowego x wyłaniany jest zwycięzca o numerze v .

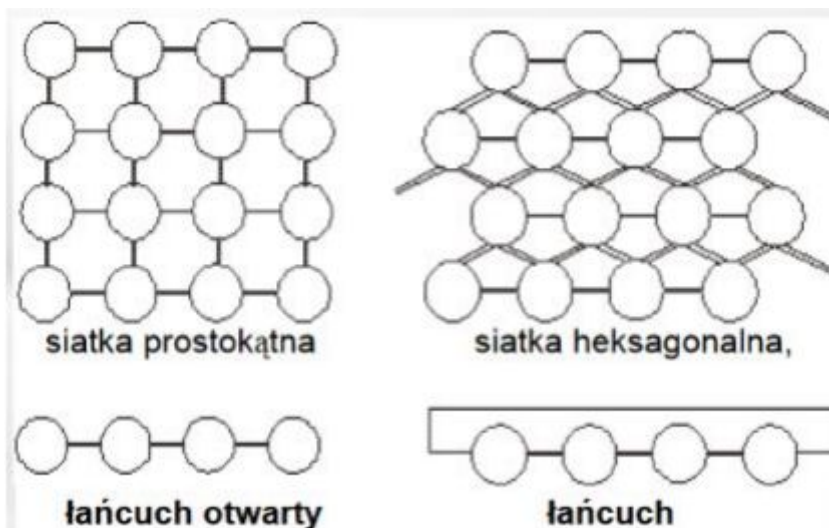
$$w_v^T x = \max_{i=1,2,\dots,k} (w_i^T x)$$

Aktualizacja wag neuronu zwycięzcy (neurony przegrywające mają na wyjściu stan 0, co blokuje proces aktualizacji ich wag). Aktualizacja ta odbywa się według reguły Kohonena:

$$w_v(t+1) = w_v(t) + \eta[x - w_v(t)]$$

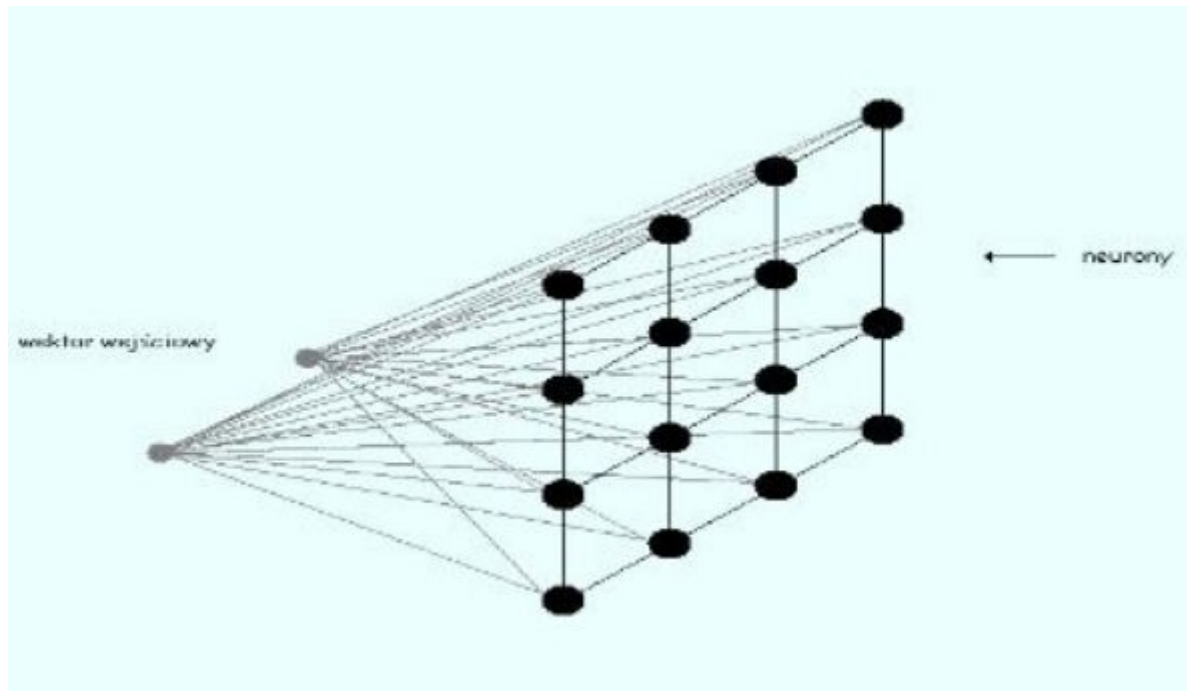
gdzie η – wsp. uczenia ($0 < \eta < 1$) i maleje w miarę postępu nauki. Wektor wag neuronu zwycięzcy jest zwiększany o ułamek różnicy $x - w$, w wyniku czego w następnych krokach lepiej odtwarza rozpatrywany wektor wejściowy.

Topologię sieci Kohonena można określić poprzez zdefiniowanie sąsiadów dla każdego neuronu. Jednostka, której odpowiedź na nasze pobudzenie jest maksymalna nazywany obrazem pobudzenia. Sieć jest uporządkowana, jeśli topologiczne relacje między sygnałami wejściowymi i ich obrazami są takie same.



Ważnym parametrem jest określenie ile neuronów obok ma podlegać uczeniu w przypadku zwycięstwa danego neuronu

Powstawanie prostokątnej dwuwymiarowej sieci Kohonena



Wykorzystaliśmy funkcję do tworzenia map samoorganizacji selforgmap. W ramach wykonywania programu otrzymaliśmy 6 różnych wykresów, które mają swoje znaczenie:

- SOM Topology – jeden kształt symbolizuje jeden neuron. Są one rozmieszczone w siatce o określonych wymiarach. Ułożenie jest nieprzypadkowe – ich sąsiedztwo może wskazywać na ich podobieństwo.
- SOM Neighbor Distances – Na tej mapie otrzymamy możliwość sprawdzenia, jak bardzo połączone są ze sobą poszczególne neurony – czyli jak silne podobieństwo one wykazują. Im jaśniejsze połączenie, tym bardziej te dane są do siebie podobne. Ciemne linie mogą zatem oznaczać granice klas.
- SOM Neighbor Connections – na tym wykresie umieszczone są połączenia między sąsiadami. Sąsiadami są zwykle próbki do siebie podobne.
- SOM Weight Planes – jest to zestaw wykresów, który wskazuje na rozkład wag poszczególnych neuronów w zależności od cechy. Im ciemniejszy kolor, tym większą wagę dany neuron skupia. Im więcej neuronów o podobnych kolorach w danym sąsiedztwie, tym te neurony są bardziej ze sobą skorelowane
- SOM Sample Hits – ten wykres pokazuje nam, ile podobnych wyników otrzymaliśmy dla danej klasy– im wyższa liczba, tym więcej obiektów o podobnych cechach można wykazać
- SOM Weight Positions – zielone kropki oznaczają dane wejściowe, a linie je łączące wskazują na korelacje pomiędzy poszczególnymi neuronami.

Przebieg ćwiczenia

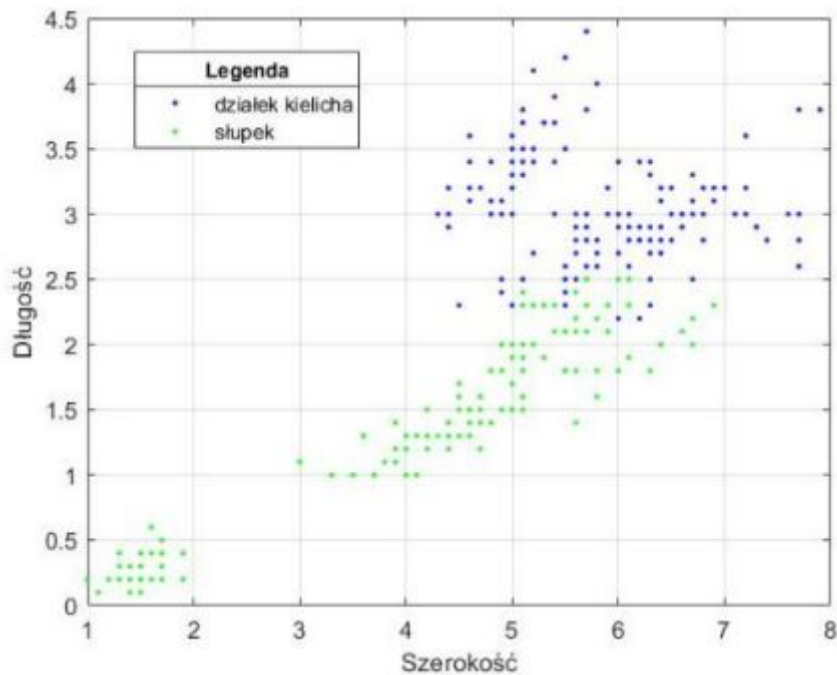
Zaimportowanie z biblioteki programu MatLab "iris_dataset", zawierającej dane na temat kwiatów typu Iris.

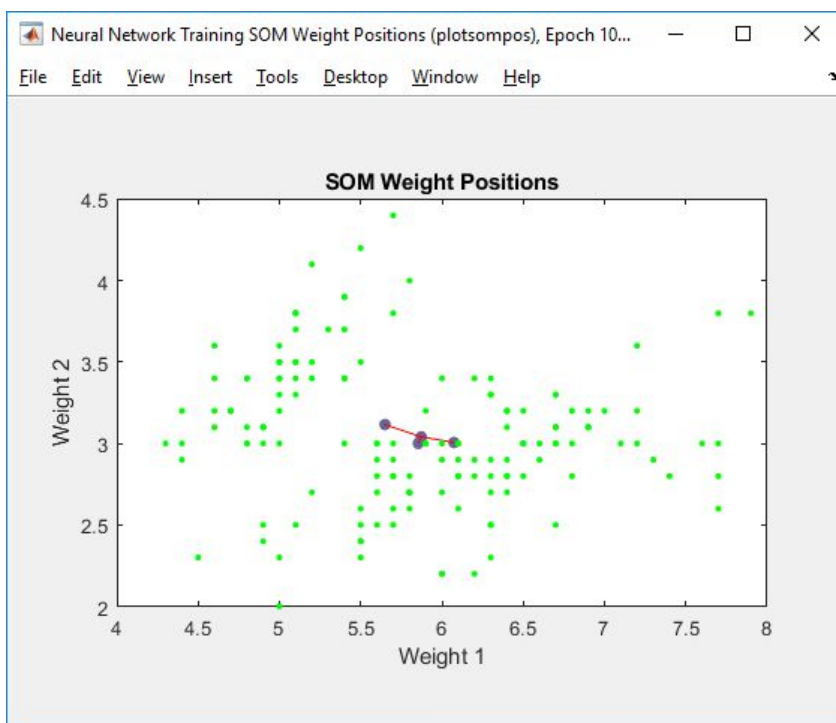
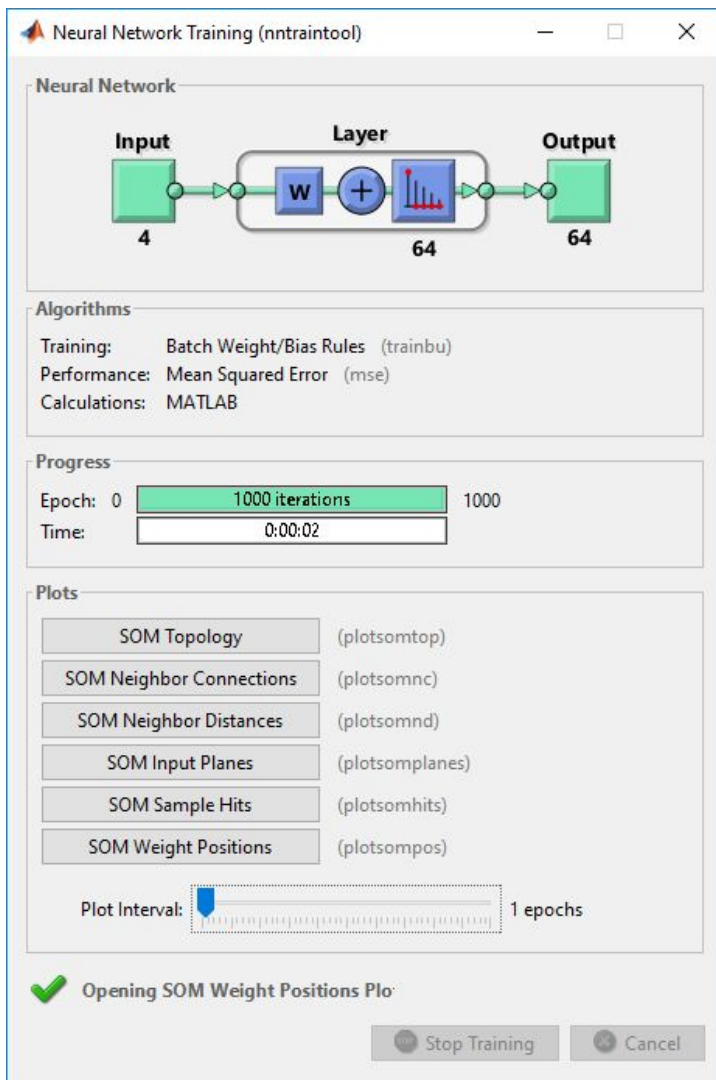
Dane te odpowiednio przedstawiają:

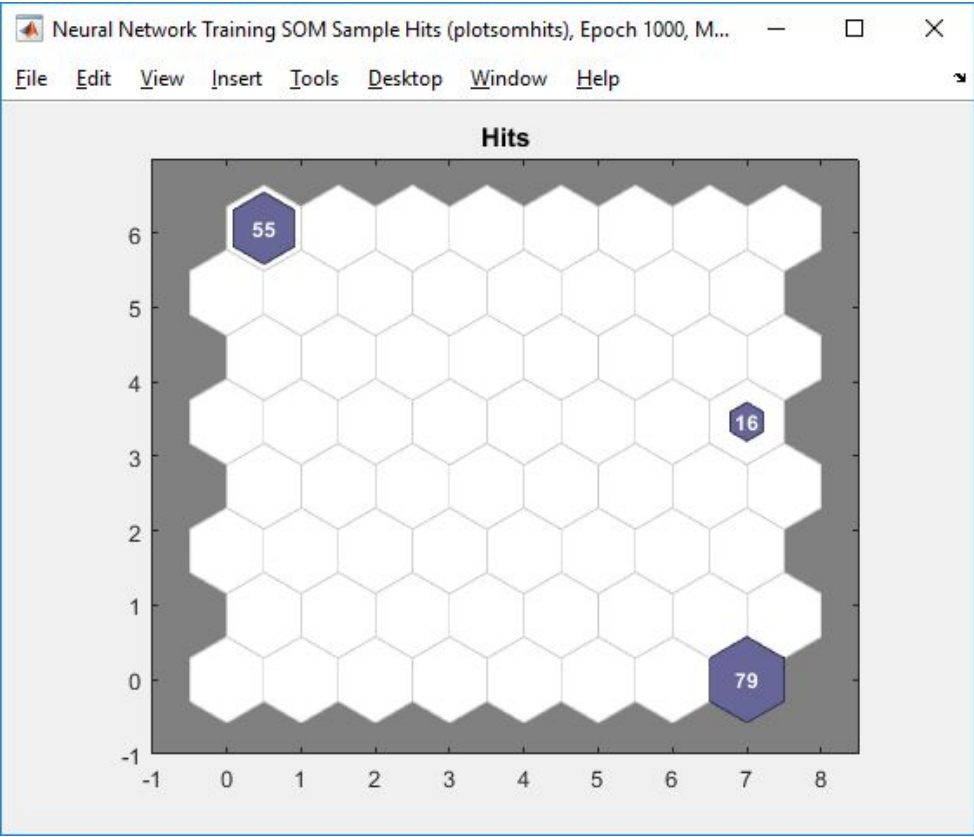
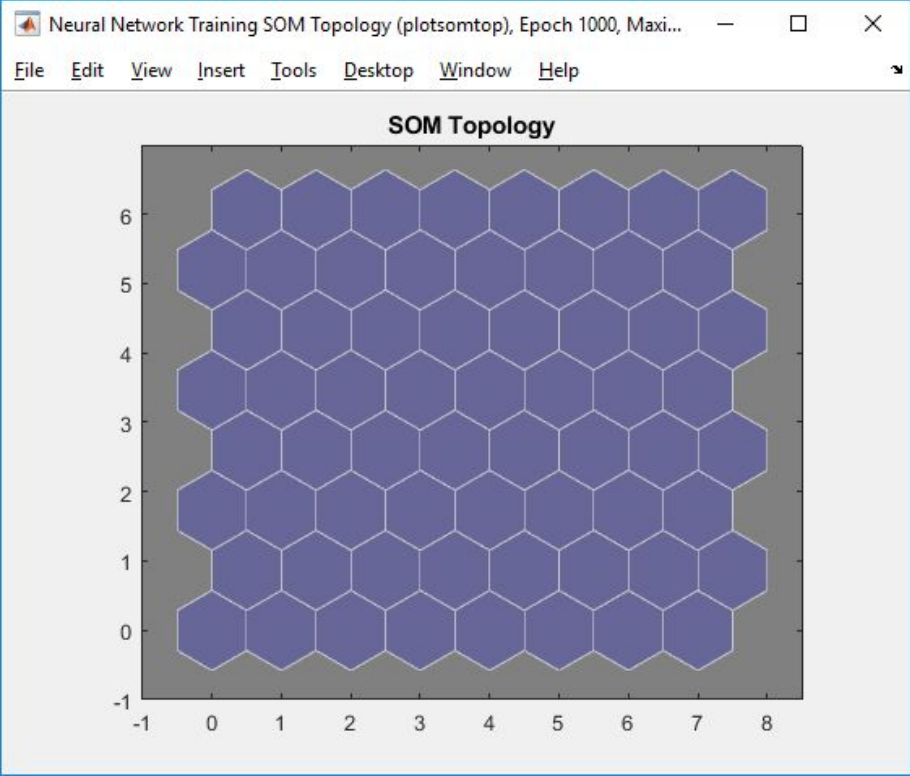
- Szerokość działki kielicha kwiatu
- Szerokość płatków kwiatowych - Wysokość działki kielicha kwiatu
- Wysokość płatków kwiatowych

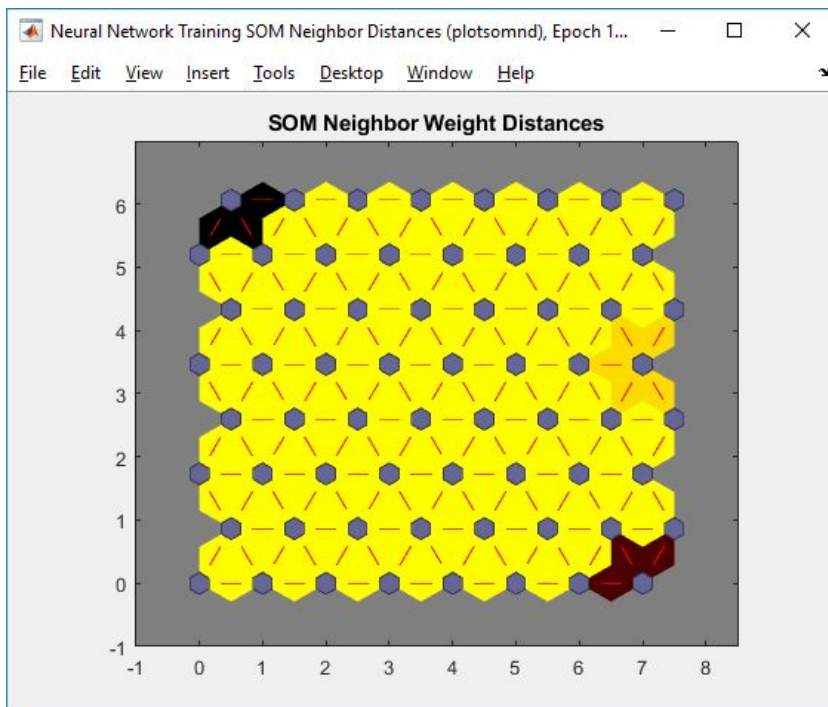
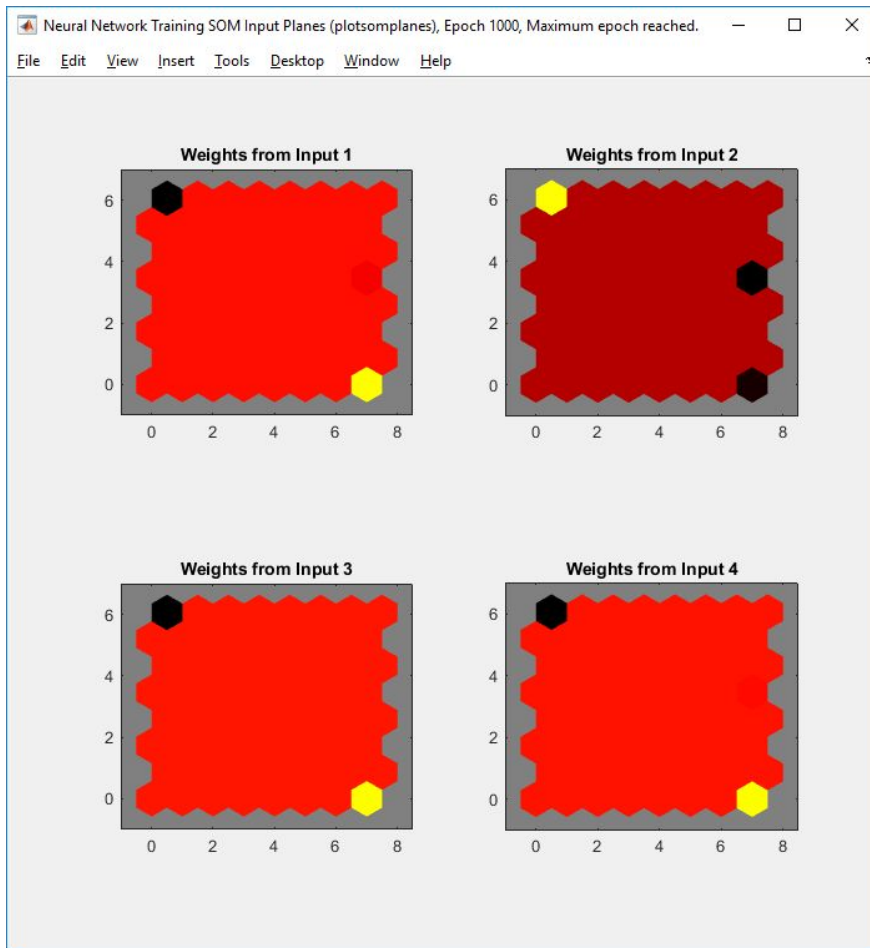
Dla trzech rodzajów kwiatów: Setosa, Virginica, Versicolor

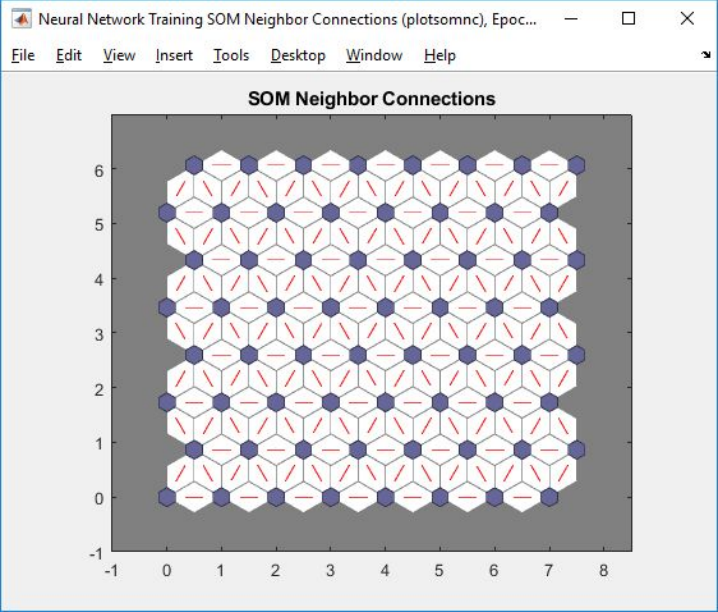
Sieć została stworzona z domyślnymi wymiarami 8 x 8 w kształcie sześciokątów, natomiast początkowy rozmiar sąsiedztwa ustawiony jest na 0, liczba epok = 1000

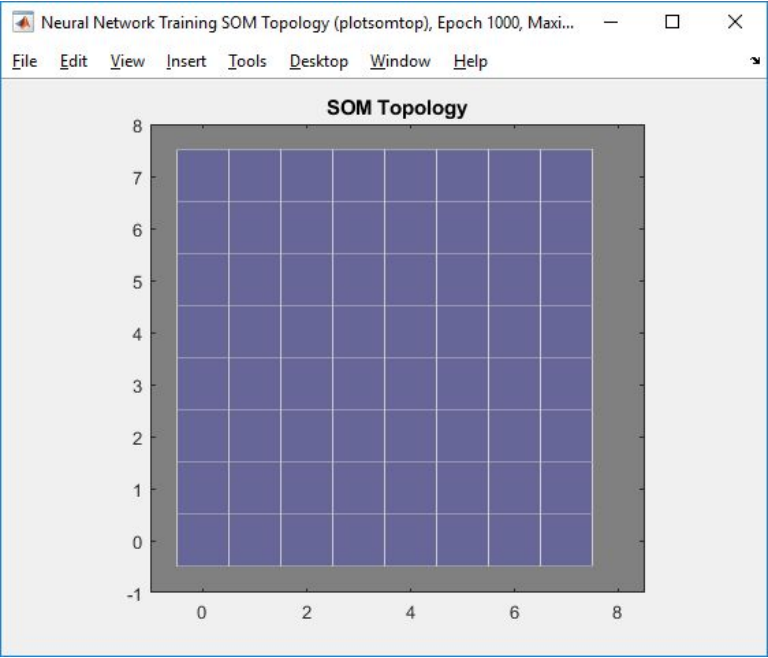


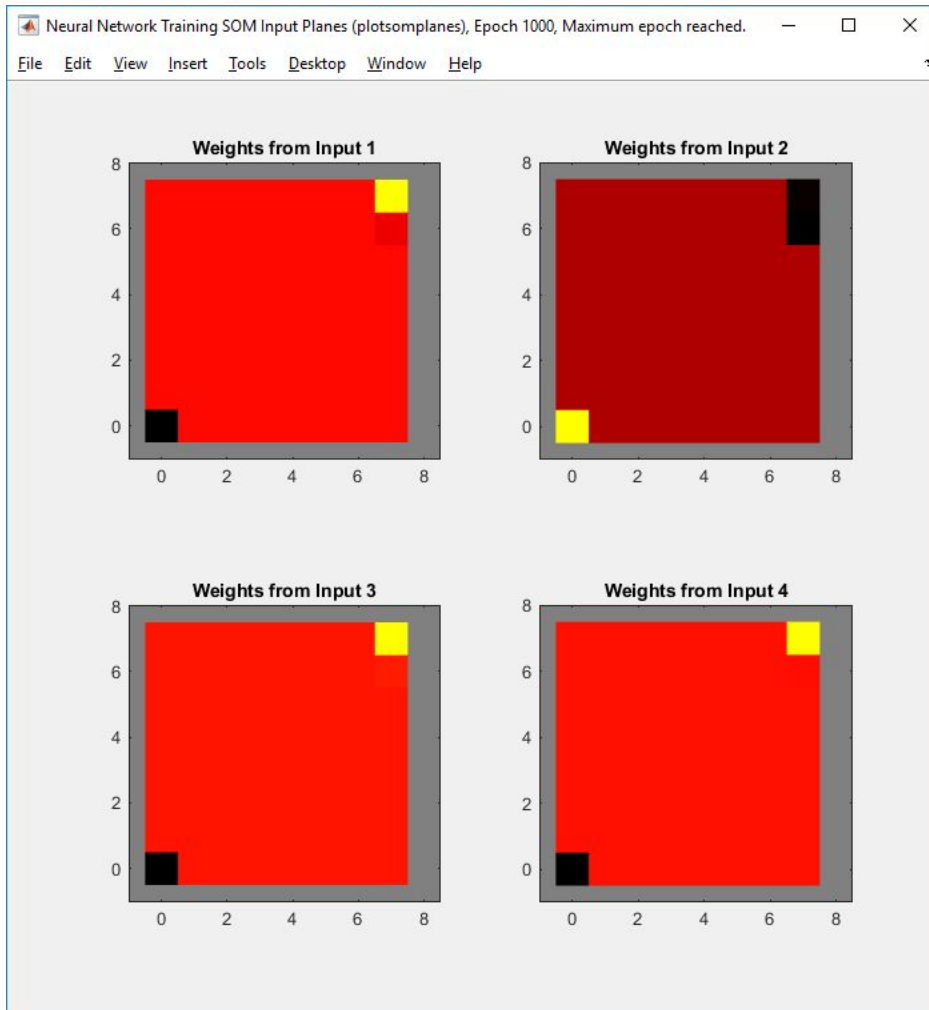
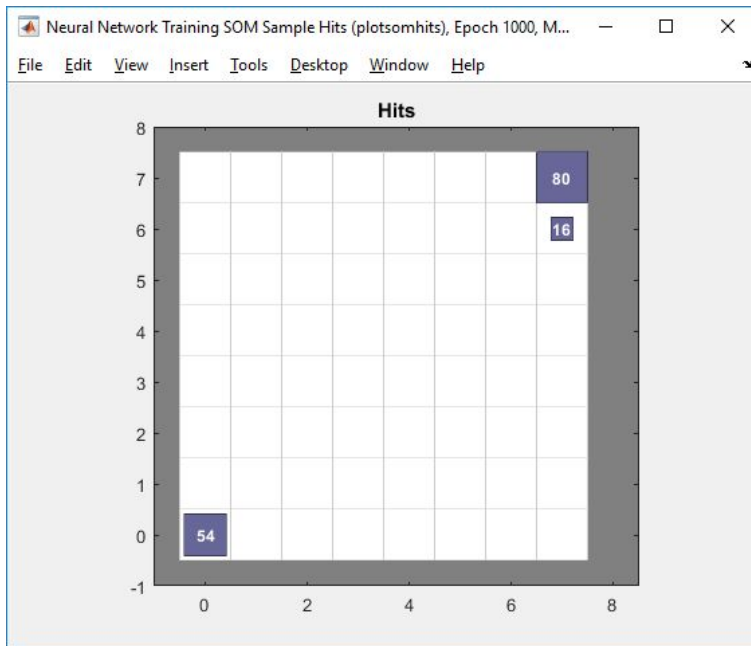


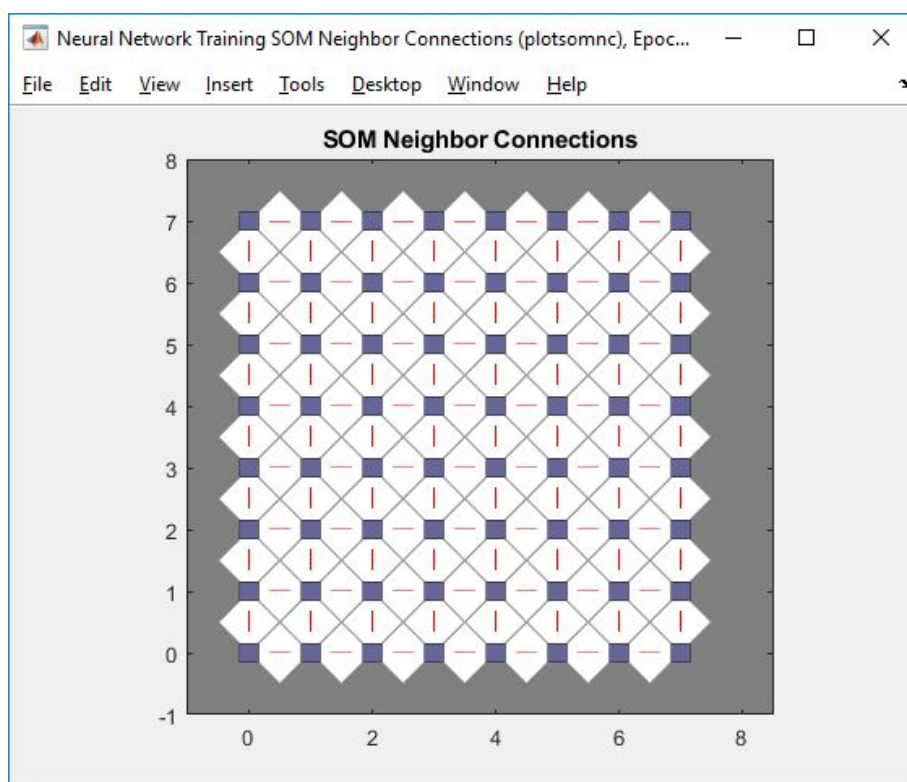
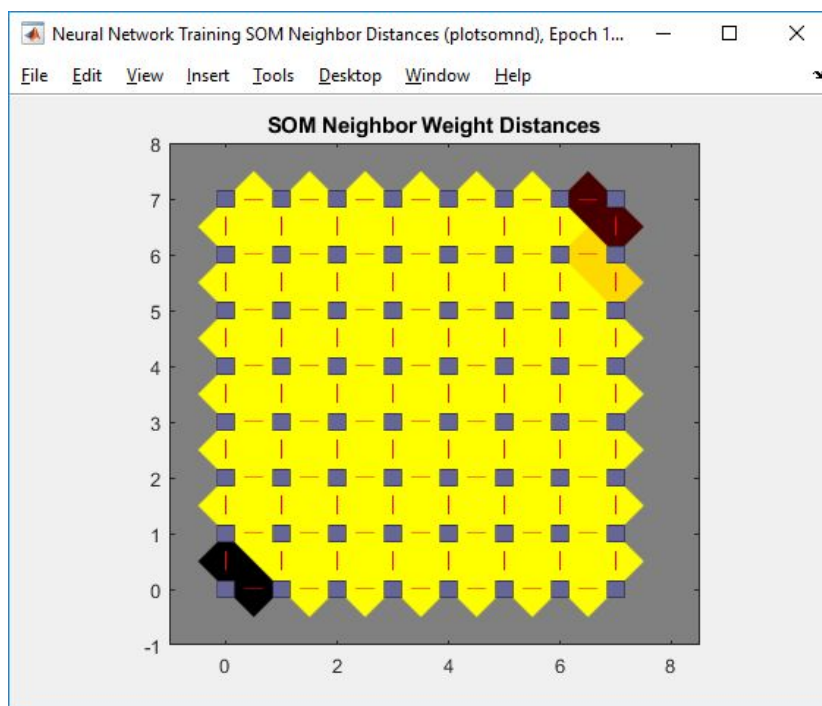












W wyniku analizy otrzymanych wyników możemy zdecydowanie stwierdzić iż topologia Hexagonalna posiada więcej połączeń pomiędzy neuronami niż topologia typu GridTop, co przy większym rozkładzie może wpłynąć na dobór odpowiednich wag dla poszczególnych neuronów. Analizując rozłożenie wag wyznaczonych przez algorytm możemy założyć iż wynikami są najmocniejsze neurony wyselekcjonowane przez WTA, niemniej jednak jeżeli zrezygnujemy z ograniczeń do zwycięskiego neuronu (WTA) i skorzystamy z samej sieci Kohonena ustalając początkowy rozmiar sąsiedztwa jako 3 otrzymamy znacznie bogatszy wysyp danych, przynajmniej dla zadanego tematu.

Wnioski

Sieci samouczące są efektywne w grupowaniu elementów i znajdowaniu ich konkretnych cech.

W algorytmie WTA sąsiedztwo nie jest istotne, za to są istotne wagi.

Im wyższy współczynnik uczenia, tym korelacja zachodzi rzadziej.

Bardziej szczegółowa jest topologia heksagonalna. Daje ona więcej informacji o podobieństwie i korelacji.