**PSI SCENARIUSZ 4**

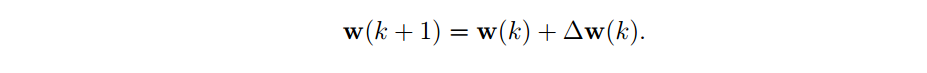
**Anna Ruchała, IS III, GR 3**

**Uczenie sieci regułą Hebba.**

W neuro-fizjologii zauważono, że waga powiązań między dwoma neuronami wzrasta przy jednoczesnym pobudzeniu oby neuronów, w przeciwnym przypadku maleje. Zaobserwowano również, że im częściej jakiś bodziec dochodzi do neuronu, tym silniejsza odpowiedź neuronu na ten bodziec. W konsekwencji najczęściej powtarzające się mają największy wpływ na samo-adaptację wag.

W sieciach neuronowych neurony ‘uczą się’ rozpoznawać bodźce (w naszym zrozumieniu: dane podawane na wejście sieci) czyli dane. Zmiany aktualnych wag powinny iść w kierunku dostosowywania wag do najczęściej przedstawianych wzorców, jakimi są wektory danych. Sygnały napływające do neuronu noszą również nazwę sygnałów presynaptycznych. Sygnały wytwarzane przez neuron to sygnały postsynaptyczne.

Podczas procesu uczenia, w miarę napływających wzorców (wektorów danych w naszym rozumieniu) sieć ‘uczy się’ tych danych, czyli adaptuje sekwencyjnie swoje wagi -aby umieć prawidłowo rozpoznawać te dane. Adaptacja ta jest iteracyjna, tzn po przedstawieniu k-tego wzorca **x(k)** wytworzone już wagi **w** zmieniają się według ogólnej zasady:



Jak wyrazić zmianę wektora wag czyli wielkość **Δw(k)**?

Zagadnienie to było przedmiotem intensywnej dyskusji, w rezultacie których pojawiło się kilka propozycji, które przetrwały do dzisiaj. W latach 1949 ukazała się książka Hebba dyskutująca to zagadnienie.

Ogólna reguła Hebba mówi, że zmiany wag powinny odbywać się według reguły:

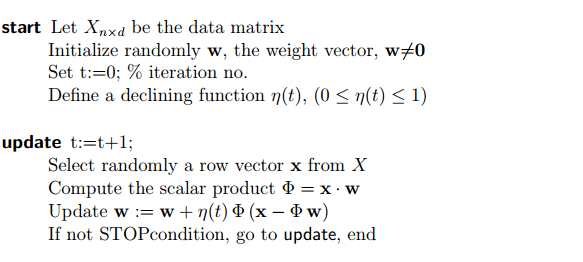


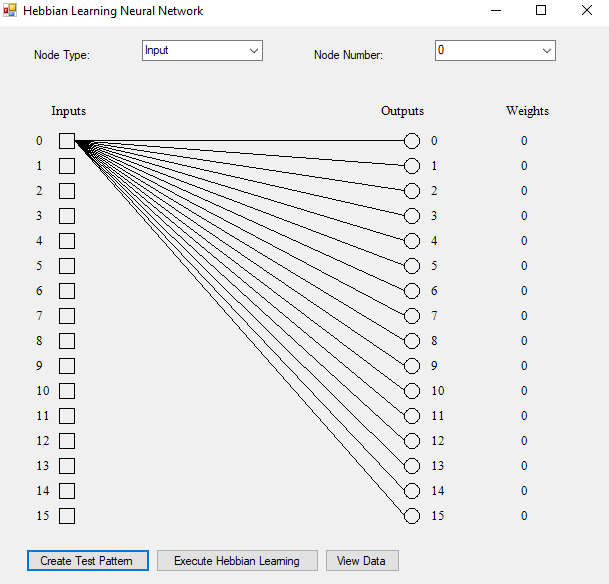
Czyli ogólnie, ze przyrost wag Δw(k) powinien zależeć zarówno od wielkości wzorca presynaptycznego x(k) jak i od wytwarzanego wzorca postsynaptycznego y(k).

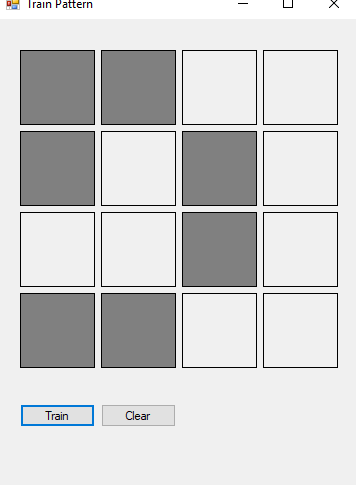
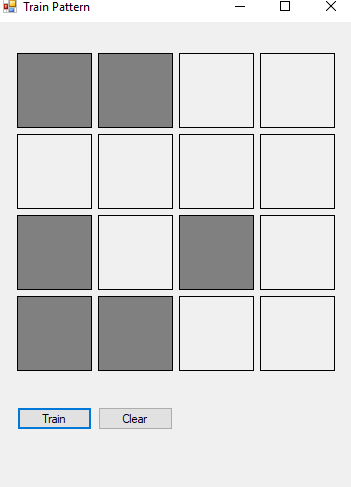
Funkcję F podaje tzw prosta reguła Hebba. Upraszcza ona F do funkcji iloczynowej:

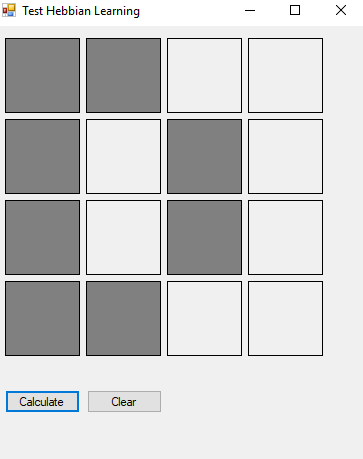


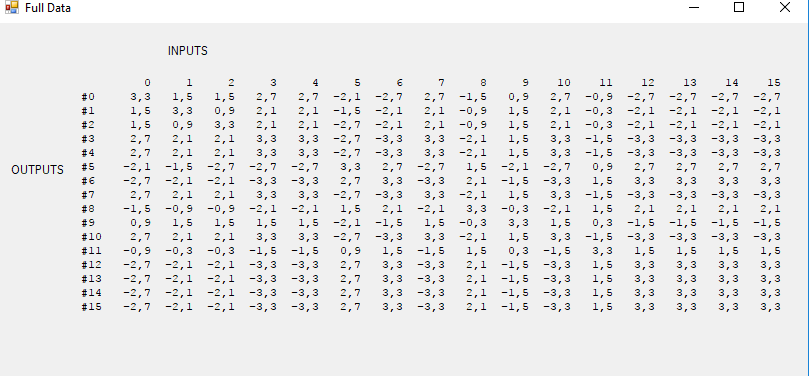
Pseudokod reguły:

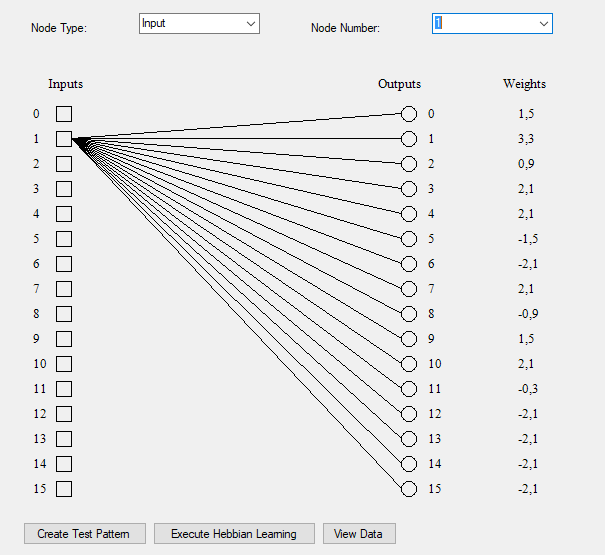


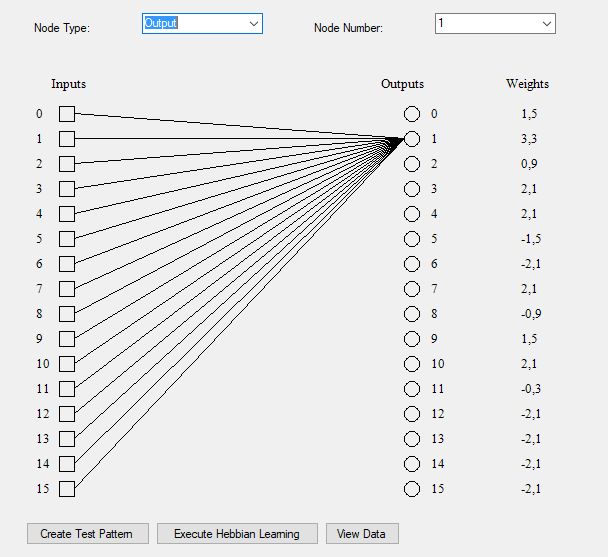












Wnioski:

Proces samouczenia ma niestety wady. W porównaniu z procesem uczenia z nauczycielem samouczenie jest zwykle znacznie powolniejsze. Co więcej bez nauczyciela nie można z góry określić, który neuron wyspecjalizuje się w rozpoznawania której klasy sygnałów. Stanowi to pewną trudność przy wykorzystywaniu i interpretacji wyników pracy sieci. Co więcej - nie można określić, czy sieć uczona w ten sposób nauczy się wszystkich prezentowanych jej wzorców. Dlatego sieć przeznaczona do samouczenia musi być większa niż sieć wykonująca to samo zadanie, ale trenowana w sposób klasyczny, z udziałem nauczyciela. - Szacunkowo sieć powinna mieć co najmniej trzykrotnie więcej elementów warstwy wyjściowej niż wynosi oczekiwana liczba różnych wzorów, które sieć ma rozpoznawać.

Bardzo istotną kwestią jest wybór początkowych wartości wag neuronów sieci przeznaczonej do samouczenia. Wartości te mają bardzo silny wpływ na ostateczne zachowanie sieci, ponieważ proces uczenia jedynie pogłębia i doskonali pewne tendencje istniejące w sieci od samego początku, przeto od jakości tych początkowych, „wrodzonych” właściwości sieci silnie zależy, do czego sieć dojdzie na końcu procesu uczenia. Nie wiedząc z góry, jakiego zadania sieć powinna się uczyć, trudno wprowadzać jakikolwiek zdeterminowany mechanizm nadawania początkowych wartości wag, jednak pozostawienie wszystkiego wyłącznie mechanizmom losowym może powodować, że sieć (zwłaszcza mała) może nie zdołać wystarczająco zróżnicować swego działania w początkowym okresie procesu uczenia i wszelkie późniejsze wysiłki, by znaleźć w strukturze sieci reprezentację dla wszystkich występujących w wejściowych sygnałach klas, mogą okazać się daremne. Można jednak wprowadzić pewien mechanizm wstępnego „rozprowadzania” wartości wag w początkowej fazie procesu uczenia. Metoda ta, zwana *convex combiation* modyfikuje początkowe wartości wag w taki sposób, by zwiększyć prawdopodobieństwo równomiernego pokrycia przez poszczególne neurony wszystkich typowych sytuacji pojawiających się w wejściowym zbiorze danych. Jeśli tylko dane pojawiające się w początkowej fazie uczenia nie będą różniły się istotnie od tych, jakie sieć będzie potem analizować i różnicować - metoda *convex combination* stworzy w sposób automatyczny dogodny punkt wyjścia do dalszego samouczenia i zapewni stosunkowo dobrą jakość nauczonej sieci w większości praktycznych zadań.