**Ćwiczenie 3 – Sieć wielowarstwowa uczona metodą propagacji wstecznej – techniki poprawiające skuteczność uczenia**

Ćwiczenie ma na celu praktyczne poznanie technik zmniejszających ryzyko przeuczenia sieci. Ma również na celu obserwację funkcji neuronów ukrytych w roli detektorów cech.

W ramach ćwiczenia należy użyć aplikacji zaimplementowanej do rozwiązania ćwiczenia 2. Przy jej pomocy należy wyuczyć sieć rozpoznawania wzorców z zapisanymi obrazami cyfr. Jako zbiór uczący można wykorzystać część zbioru MNIST (do pobrania z BOARD’a Łukasza Olecha). Sieć powinna być wykorzystana w roli autokodera, tzn. powinna posiadać tyle samo neuronów wejściowych, co wyjściowych, gdyż staramy się wyuczyć ją replikowania obrazu wejściowego.

Jako funkcję kosztu, proszę użyć błąd średniokwadratowy.

I. Należy sprawdzić jak wpływa na powstawanie przeuczenia:

1. Zastosowanie w trakcie uczenia regularyzacji L2. Należy porównać wyniki osiągane z regularyzacją i bez regularyzacji dla tych samych wartości parametrów i tej samej architektury.

Zmiana wag przy uczeniu po wzorcu odbywa się wówczas zgodnie ze wzorem:= 

gdzie λ’ jest współczynnikiem regularyzacji, który w swej wartości powinien uwzględniać liczebność zbioru uczącego, np. dla małej liczby wzorców przyjmiemy *0.1* natomiast dla dużej liczby, np. dla wszystkich wzorców z MNIST (jeśli ktoś zdąży ☺) można przyjąć *5.0.*

Dla uczenia paczkami (ang. batch training) np. paczka – *10* wzorców, zmiana wag według wzoru



Suma po wszystkich wzorcach w paczce

1. Dropout

Należy porównać wyniki osiągane przez sieć o liczbie neuronów w warstwie ukrytej będącej średnia arytmetyczną liczby neuronów wejściowych i wyjściowych z wynikami sieci o podwójnej liczbie neuronów w warstwie ukrytej w stosunku do poprzedniej ale takiej w której zastosowano droput.

1. Powiększanie liczby wzorców trenujących.

II. Należy wykonać wizualizację cech na jakie wrażliwe są poszczególne neurony ukryte dla sieci o najlepszej dokładności rozpoznawania dla wzorców testowych. Wizualizację wykonujemy dla każdego *j-tego* neuronu w warstwie ukrytej. W ten sposób możemy dowiedzieć się jaką cechę wykrywa dany neuron ukryty. Proszę w sprawozdaniu skomentować uzyskane wizualizacje.



Proszę przeskalować wagi w warstwie ukrytej do przedziału [0;1] lub inny dogodny do wyświetlania w skali szarości. Następnie zakładając, że wejście było wzorcem będącym obrazem cyfry *10x10* (tzn. N= 100 wejść), dla pojedynczego *j-tego* neuronu w warstwie ukrytej proszę obliczyć wartość intensywności piksela według wzoru:

Gdzie *i* to indeks *i-tego* neuronu wejściowego, czyli dla *j-tego* neuronu obliczymy *N* intensywności pikseli, które układamy w obraz o takim samym wymiarze jaki był podawany na wejście (tutaj w przykładzie *N=10x10*, dla MNIST *28x28*, dla poprzedniego ćwiczenia *7x10*). Obrazy te wyświetlić dla zwykłej sieci oraz dla takiej, w której zastosowano dropout oraz dla sieci z regularyzacją L2.