# Wstęp do Sztucznej Inteligencji – Ćwiczenie 5 – Sztuczne sieci neuronowe

## Obserwacje

Z racji na bardzo długi czas uczenia się sieci neuronowej wykorzystującej metodę stochastycznego najszybszego spadku, zdecydowaliśmy się wykorzystać jego zmodyfikowaną wersję z mini-pakietami (minibatches SGD).

Poniżej przedstawione są wykresy błędu uczonej sieci neuronowej w kolejnych epokach dla kilku wartości parametru alfa (learning rate)

Graphical user interface

Description automatically generated

Do oceny jakości sieci neuronowej zamieszczamy również wykres zmiany błędu uczonej sieci neuronowej dla zbioru danych Iris:

Chart

Description automatically generated

## Wnioski

Po pierwsze, zbyt mała wartość parametru alfa może powodować pogorszenie efektów uczenia się sieci neuronowej (alfa=0,0001 vs alfa=0,0105). To samo dotyczy zbyt dużej wartości parametru alfa (np. alfa=0,2), gdzie błąd na koniec kolejnych epok nie zmieniał się wcale lub zmieniał się marginalnie, a finalna jakość modeu była znacząco gorsza niż dla alfa równego 0,006 lub 0,0105. Skrajnym przypadkiem obrazującym problemy przy zbyt dużej wartości parametru alfa jest przypadek 5. z zamieszczonych wyżej wykresów (alfa=0,08), gdzie błąd na koniec kolejnych epok oscylował wokół prostej – taki model nie byłby użyteczny.

Po drugie, w przetestowanym i przedstwawionym na wykesie zakresie 0-800 epok, większa liczba epok poprawia jakość sieci. Sytuacja ta ma miejsce z nielicznymi wyjątkami. Jeden z nich to opisana wyżej sytuacja oscylacji wokół prostej, która sygnalizuje zły dobór parametru uczenia się. Druga sytuacja ma miejsce w teście nr 3, gdzie sieć osiągnęła najniższy błąd na koniec epoki ~120, potem sytuacja nieznacznie się pogorszyła. Takiemu zachowaniu można zapobiec poprzez dodanie mechanizmu niepozwalającego na zmianę wag sieci, jeśli ta zmiania miałaby doprowadzić do pogorszenia efektywności modelu. Należy się wtedy jednak zastanowić, czy model nie utknie wtedy w minimum lokalnym (taka sytuacja jest możliwa dla skomplikowanych funkcji kosztu).

Po trzecie, dla parametru alfa=0,0105 procent dobrze przewidzianych klas czerwonego wina jest bardzo zbliżony do skuteczności klasyfikatora z poprzedniego ćwiczenia (oba modele ok 56% skuteczności). Różnica jest na tyle nieduża, że można uznać, że przy innym wylosowaniu kolejności danych w zbiorze treningowym róznica ta, mogłaby zostać zniwelowana.

Po czwarte, z wykresu zmiany błędu uczonej sieci neuronowej dla zbioru danych Iris widać, że algortym uczenia sieci neuronowej został zaimplementowany właściwie (otrzymany wynik to 96.67 % skuteczności).