

## Laboratorium 5 – sztuczne sieci neuronowe

## Semestr 21Z

Gierulski Wojciech, Kaniuka Jan

## Treść zadania

Zaimplementować sztuczną sieć neuronową z warstwą ukrytą. Implementacja powinna być elastyczna na tyle, żeby można było zdefiniować liczbę neuronów wejściowych, ukrytych i wyjściowych. Wykorzystać sigmoidalną funkcję aktywacji i do trenowania użyć wstecznej propagacji błędów z użyciem metody stochastycznego najszybszego spadku. Sieć nauczyć rozpoznawać jakość wina jak w ćwiczeniu 4. i porównać wyniki z otrzymanymi poprzednio. Na wykresie pokazać jak zmieniał się błąd uczonej sieci w kolejnych epokach. Poeksperymentować ze współczynnikiem uczenia oraz liczbą epok.

## Założenia wstępne

Przyjęto następującą postać funkcji straty (**loss function**):

$$q(y) = \frac{1}{2} ||y - y^d||^2$$

gdzie  $y$  oznacza wektor wyjść sieci, a  $y^d$  wektor wyjść oczekiwanych.

Początkowo **wagi sieci zainicjalizowano losowymi wartościami** z przedziału (0,1), a dane wejściowe **znormalizowano**.

Liczbę neuronów warstwy ukrytej wyznaczono wg źródeł jako tzw. *rule of thumb* na podstawie wzoru

$$\frac{2}{3} * \text{liczba neuronów wejściowych} + \text{liczba neuronów wyjściowych} = 19$$

Następnie sprawdzono, że taka wartość daje zadowalające rezultaty.

Do oceny sieci zastosowano **walidację krzyżową** 3-krotną.

Jako metrykę oceny jakości zastosowano **accuracy** =  $\frac{\text{liczba poprawnych klasyfikacji}}{\text{liczba wszystkich klasyfikacji}} \times 100\%$

Eksperymenty prowadzono na zbiorze **winequality-red.csv**.

## Wyniki eksperymentów

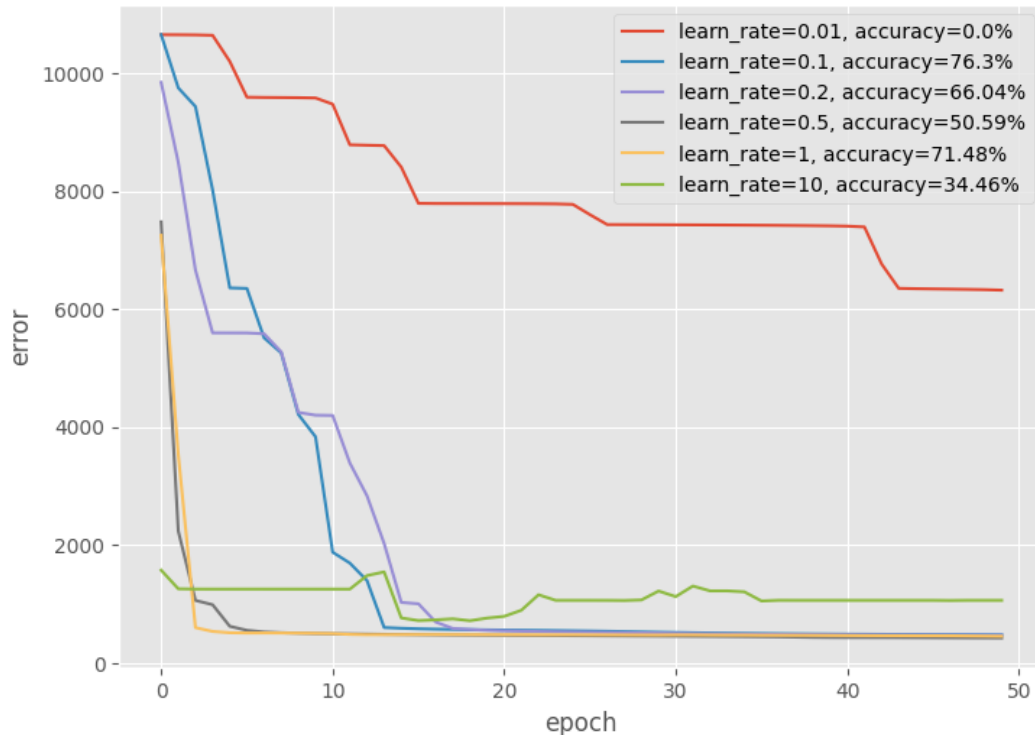
Na rysunku 1 przedstawiono, jak zmieniał się błąd uczonej sieci w kolejnych epokach, dla różnych współczynników uczenia. Jeżeli **learn rate** jest zbyt mały to uczenie trwa długo i przy niewystarczającej liczbie epok jej działanie jest dalekie od pożądanego. Zbyt duża wartość **learn rate** może spowodować „przeskakiwanie” rozwiązania zapewniającego minimalną wartość funkcji błędów. Z eksperymentów wynika, że najlepszy model otrzymujemy dla współczynnika uczenia równego 0,1.

W tabeli 1 przedstawiono dokładność modelu dla różnych liczb epok. Przyjęto stały współczynnik uczenia równy 0,2. Z tabeli wynika, że najlepsze rezultaty uzyskujemy dla liczby epok z zakresu 30-60.

Liczba epok	5	10	20	30	40	50	60	80	100	150	200
accuracy [%]	0	67,97	65,10	75,35	74,17	68,60	75,60	58,72	72,29	52,28	64,10

Tabela 1

### Błąd uczonej sieci w kolejnych epokach dla różnych współczynników uczenia



Rys. 1

### Porównanie z naiwnym klasyfikatorem Bayesa

Najlepsze uzyskane *accuracy* za pomocą klasyfikatora Bayesa wyniosło **56%**. Zastosowanie sztucznej sieci neuronowej do klasyfikacji pozwoliło uzyskać wynik na poziomie **77%**. Jest to duży wzrost jakości klasyfikatora, jednak wciąż pozostaje on daleki od bycia skutecznym. Powodem tego może być słaba jakość danych (brak danych dla niektórych klas jakości win, nierównomierny rozkład jakości win).

### Obserwacje

W przypadku sztucznej sieci neuronowej możemy dostrajać wartości wielu parametrów (np. liczba neuronów, liczba warstw, metoda trenowania, współczynnik uczenia, rodzaj funkcji aktywacji), gdzie w przypadku naiwnego klasyfikatora Bayesa nie było takiej możliwości. Potencjalnie daje nam to możliwość uzyskania lepszych rezultatów, jednak potrzebne do tego eksperymenty będą z pewnością czasochłonne.