

Mierzę:

- czas predykcji
- skuteczność sieci na danym zbiorze danych (dla dwóch zbiorów danych)

Zmieniam:

- liczbę warstw (tyle ile jest teraz, +1, -1)
- liczba jednostek w warstwie (tyle co teraz, o 0.5 mniej, o 0.5 więcej w środkowych warstwach)
- funkcja aktywacji (relu, tanh)
- optymalizator (Adam, SGD)
- współczynnik uczenia (0.001, 0.0001, 0.00001)
- batch_size (4, 16, 32)

Model podstawowy:

- liczba warstw w pełni połączonych - 4
- suma liczby jednostek w warstwie - (184, 212, 106, 4) -> pierwszy zbiór, (248, 340, 170, 9) -> drugi zbiór
- funkcja aktywacji - relu
- optymalizator – Adam
- współczynnik uczenia – 0.001
- batch_size - 4

Podczas badań zmieniam jeden parametr na raz i liczę metryki. Po każdej zmianie modelu liczę metryki po 10? razy dla każdego zbioru I wyciągam średnią z 10 pomiarów.

Zabawa danymi na zbiorze ze zdjęciami miejsc i danymi audio (104 kolumny).

Do badań biorę:

- 10 kolumn o największej korelacji z klasą wynikową
- 10 kolumn o najmniejszej korelacji
- 20 kolumn o korelacji najbliższej średniej z wszystkich korelacji

Porównuję wyniki dla różnej liczby kolumn tylko na modelu podstawowym.