

Sieci neuronowe

Multi-layer Perceptron (MLP)

Paulina Żak

Różnice w przeprowadzonym preprocessingu danych miały bardzo duży wpływ na wynik końcowy dla dokładności predykcji.

W raporcie zostały przedstawione punkty dla zestawu danych z oraz bez standaryzacji

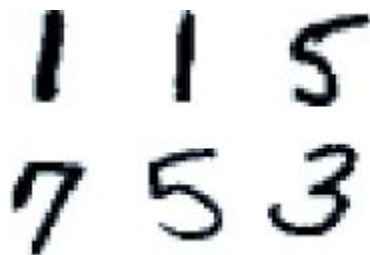
```
X_train = preprocessing.scale(X_train)
X_test=preprocessing.scale(X_test)
```

Do utworzenia modeli wykorzystałam klasy z

`sklearn.neural_network` oraz `sklearn.neighbors`:

- `MLPClassifier`
- `KNeighborsClassifier`
- `NearestCentroid`

MNIST



MNIST to dataset zbudowany z ręcznie pisanych cyfr jako obrazy 28x28

- Liczba cech 784
- Zestaw treningowy 60'000
- Zestaw testowy 10'000

1) Model 1 warstwy

- algorytm : stochastic gradient descent
- warstwy ukryte : 50
- α : 0.0001
- maksymalna liczba iteracji : 30

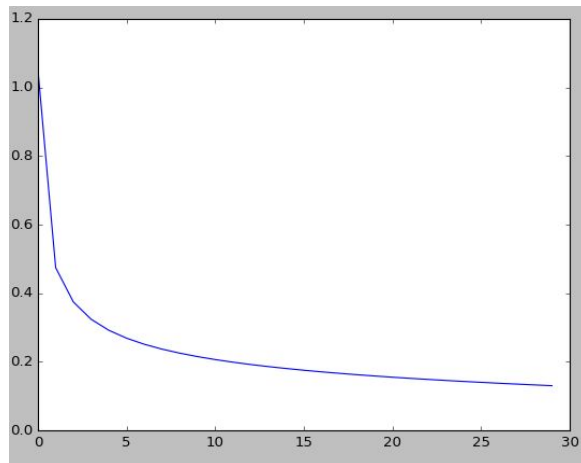
Accuracy score:

ze standaryzacją 0.9547

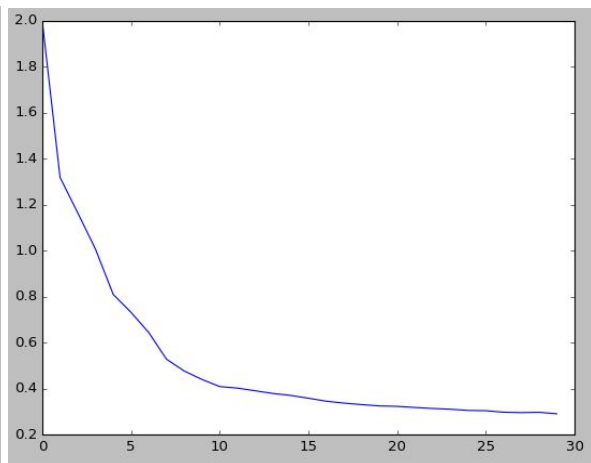
bez standaryzacji 0.8697

Loss function:

ze standaryzacją



bez standaryzacji



2) 3 modele

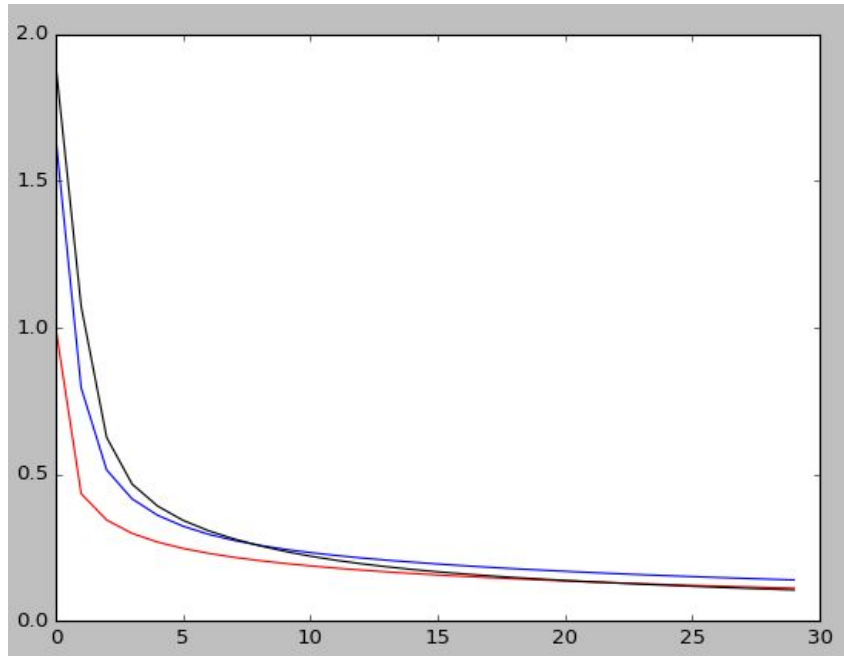
- algorytm : stochastic gradient descent
- warstwy ukryte :
 - czerwony 100
 - niebieski 30,10
 - czarny 50,30,10
- α : 0.0001
- maksymalna liczba iteracji : 30

Accuracy scores:

- warstwy 100 : ze standaryzacją 0.9579, bez standaryzacji 0.9222
- warstwy 30,10: ze standaryzacją 0.9464, bez standaryzacji 0.1135
- warstwy 50,30,10: ze standaryzacją 0.9562, bez standaryzacji 0.1135

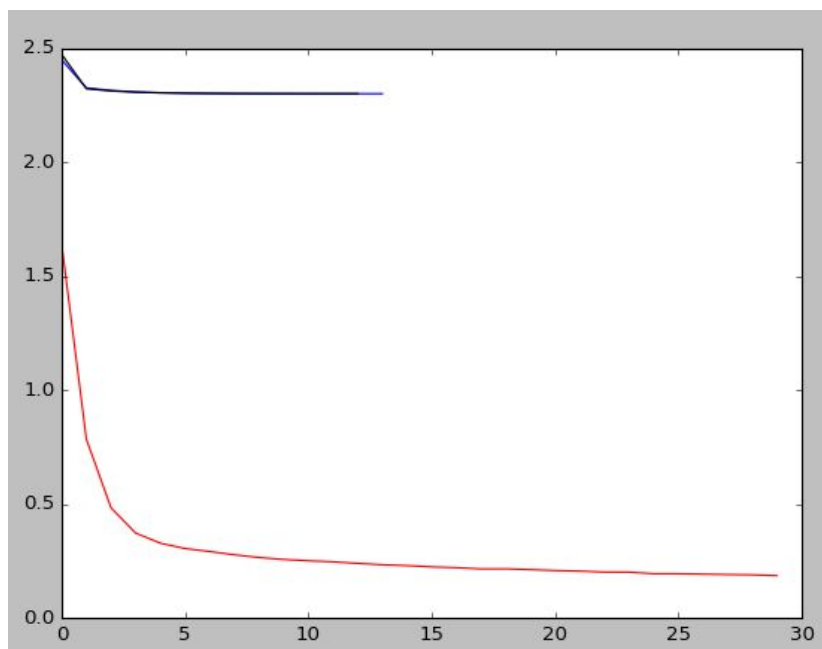
Wyniki loss function wraz ze standaryzacją:

Jak widać na obrazku poniżej mimo tego, że sieć o mniejszej ilości warstw najszybciej się nauczyła, to od mniej więcej od 15 iteracji wyniki dla wszystkich trzech modeli były porównywalne.



Loss function bez standaryzacji

Nauka bez standaryzowanych danych dla wielu warstw zatrzymała się już po pierwszej iteracji.



3) Model LBFGS

- algorytm : LBFGS
- warstwy ukryte : 50
- α : 0.0001
- maksymalna liczba iteracji : 30

Accuracy score:

ze standaryzacją 0.9387

bez standaryzacji 0.3582

Najbliższy sąsiad

Accuracy score:

ze standaryzacją 0.9691

bez standaryzacji 0.9427

Najbliższa centroida

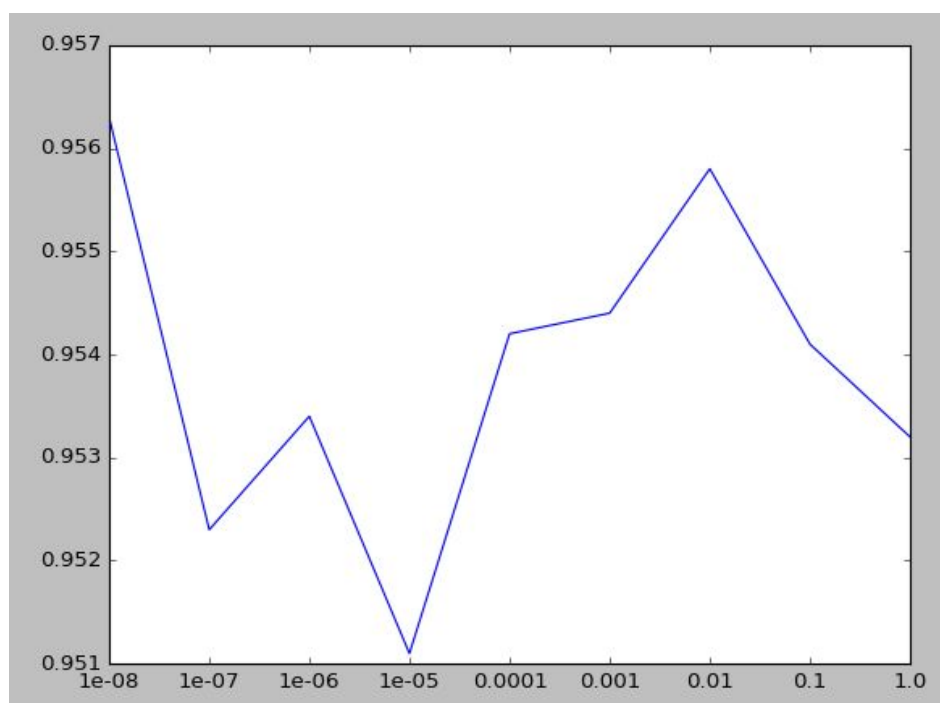
Accuracy score:

ze standaryzacją 0.8089

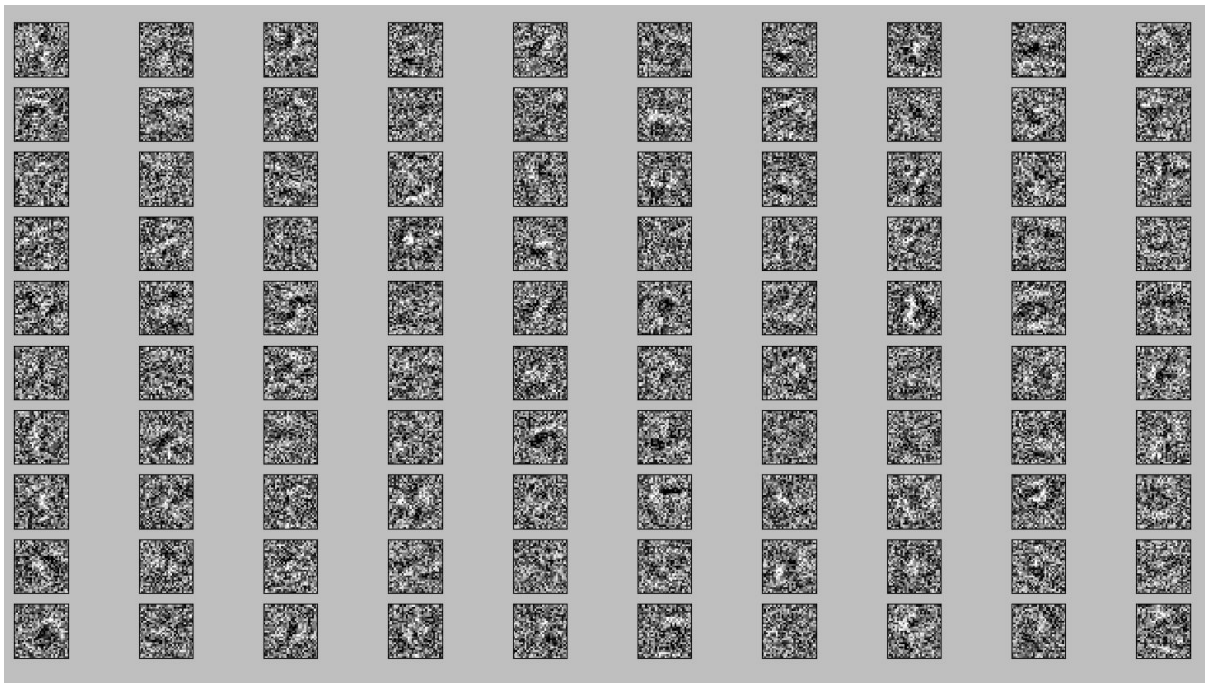
bez standaryzacji 0.8203

Wynik dokładności w zależności do alpha.

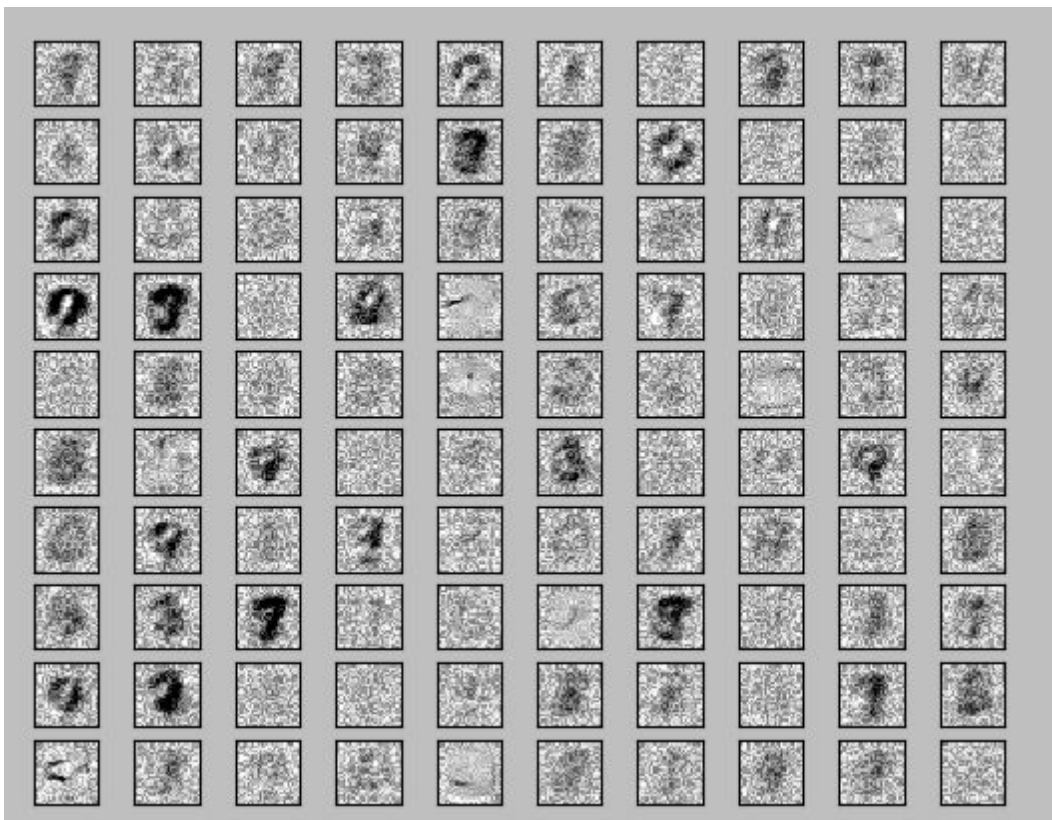
Zmiany w alpie nie wniosły żadnych znaczących różnic



Wagi MLP w formie macierzy pikseli dla datasetu standaryzowanego

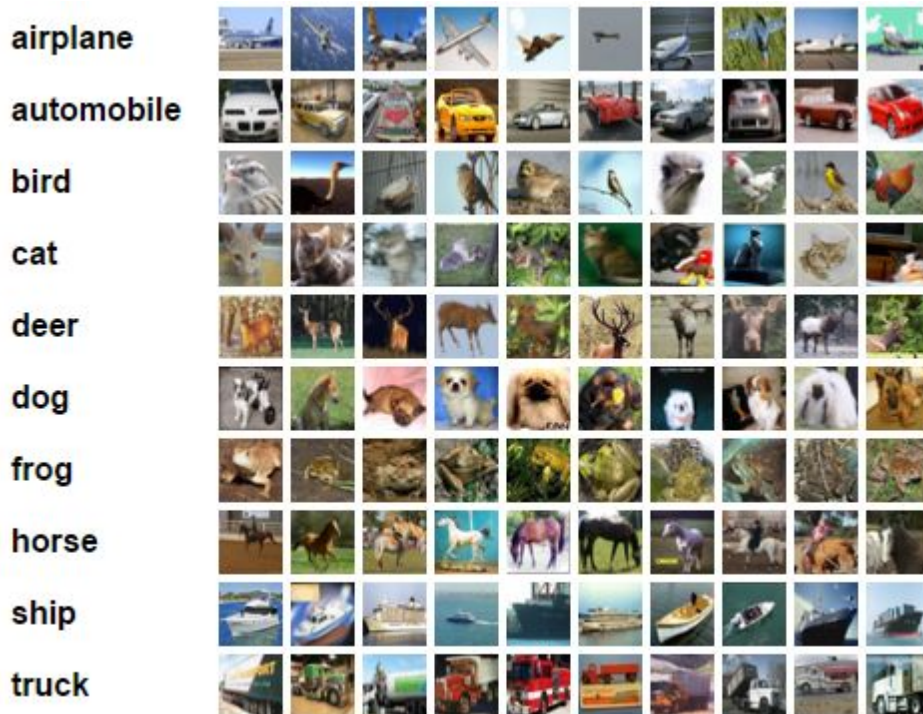


datasetu bez standaryzacji



CIFAR10

Cifar-10 jest zestawem różnych obrazów ukazujące 10 klas.



- Liczba cech: 3072
 - gdzie pierwsze 1024 to zapis koloru czerwonego
 - drugie 1024 zielonego
 - trzecie 1024 niebieskiego
- a każde 1024 opisuje obrazek 32x32
- Zestaw treningowy 50'000 (ograniczony do 10'000) ze względu na czas
- Zestaw testowy 10'000
- 1. Model 1 warstwy**
 - algorytm : stochastic gradient descent
 - warstwy ukryte : 50
 - α : 0.0001
 - maksymalna liczba iteracji : 30

Accuracy score:

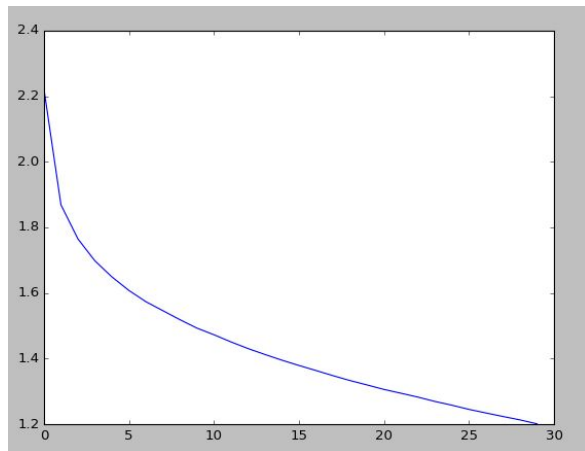
ze standaryzacją 0.4294

na pełnym zestawie 50000 0.4925

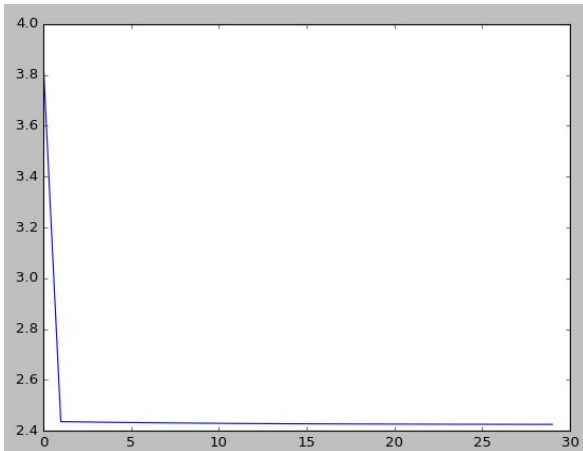
bez standaryzacji 0.1

Loss function:

ze standaryzacją



bez standaryzacji



2) 3 modele

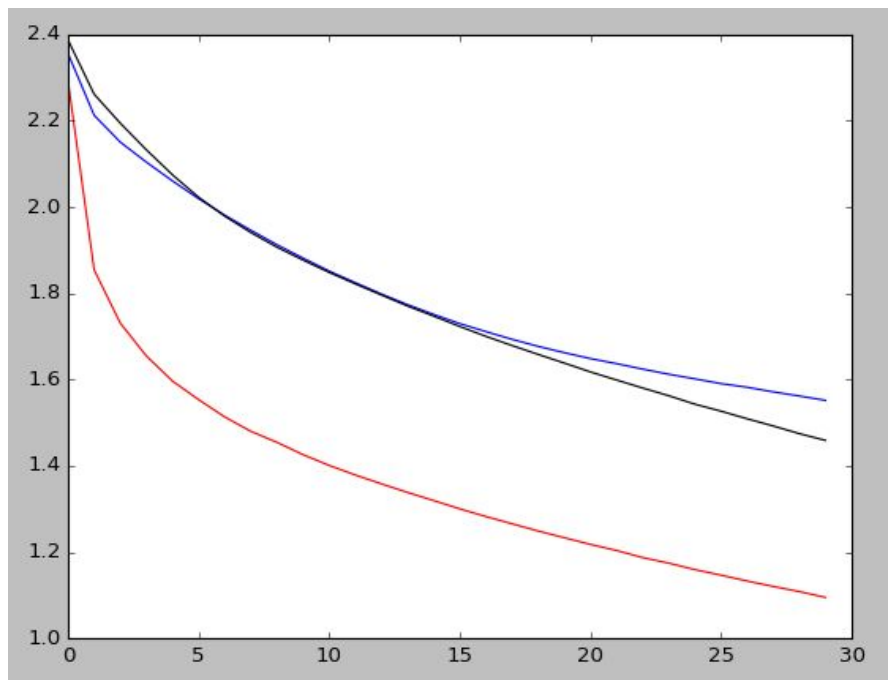
- algorytm : stochastic gradient descent
- warstwy ukryte :
 - czerwony 100
 - niebieski 30,10
 - czarny 50,30,10
- α : 0.0001
- maksymalna liczba iteracji : 30

Accuracy scores:

- warstwy 100 : ze standaryzacją 0.4337, bez standaryzacji 0.1
- warstwy 30,10: ze standaryzacją 0.3737, bez standaryzacji 0.1
- warstwy 50,30,10: ze standaryzacją 0.3897, bez standaryzacji 0.1

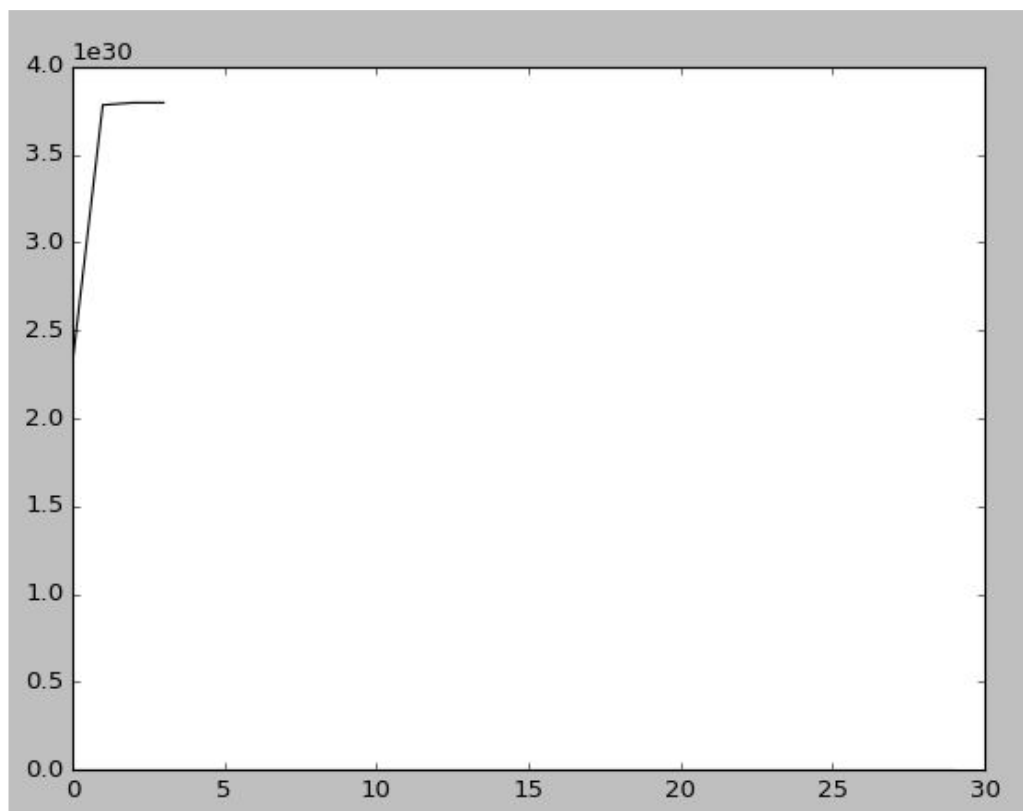
Wyniki loss function wraz ze standaryzacją:

Najlepiej nauczył się model 1 warstwy, a w dwu i trzywarstwowym praktycznie nie widać różnicy w ich nauczaniu.



Loss function bez standaryzacji

Żaden z modeli nie był w stanie nauczyć się na danych bez standaryzacji.



3) Model LBFGS

- algorytm : LBFGS
- warstwy ukryte : 50
- α : 0.0001
- maksymalna liczba iteracji : 30

Accuracy score:

ze standaryzacją 0.3777

bez standaryzacji 0.3582

Najbliższy sąsiad

Accuracy score:

ze standaryzacją 0.2908

bez standaryzacji 0.2877

Najbliższa centroida

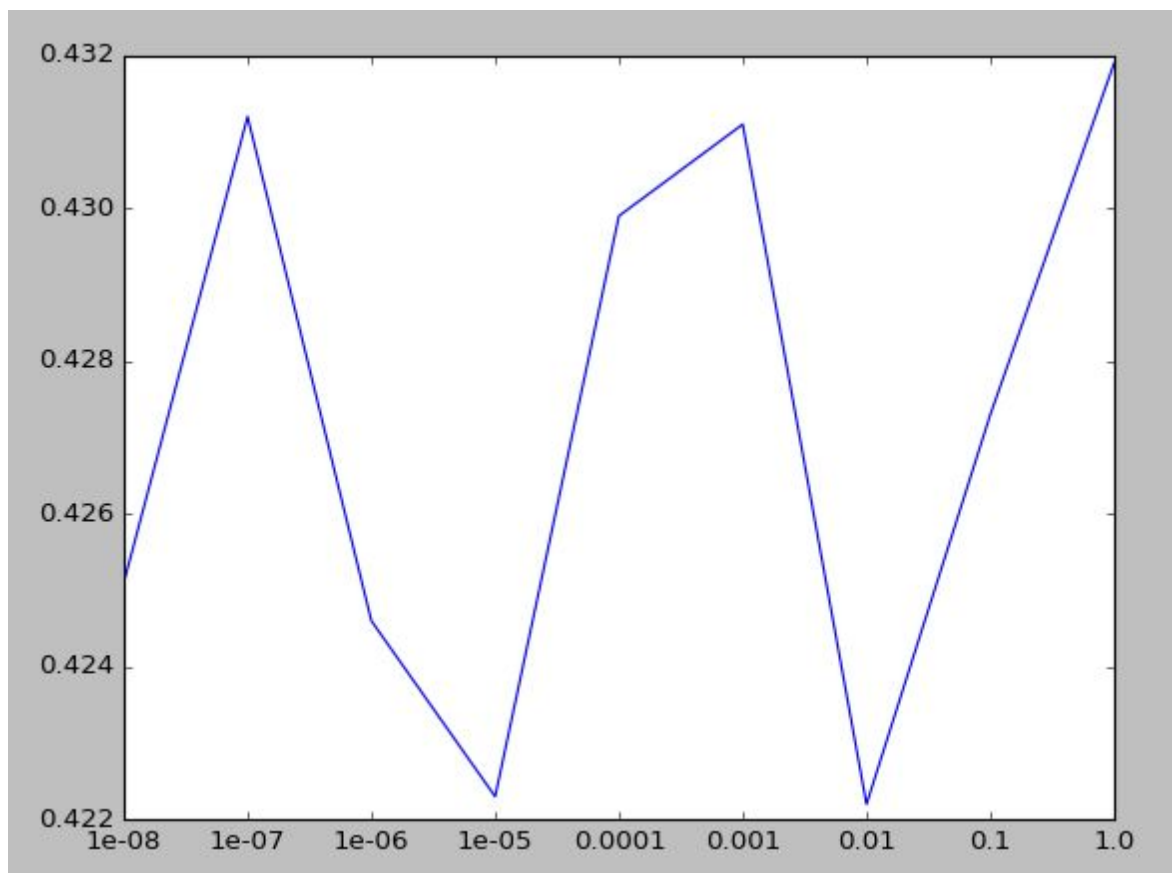
Accuracy score:

ze standaryzacją 0.278

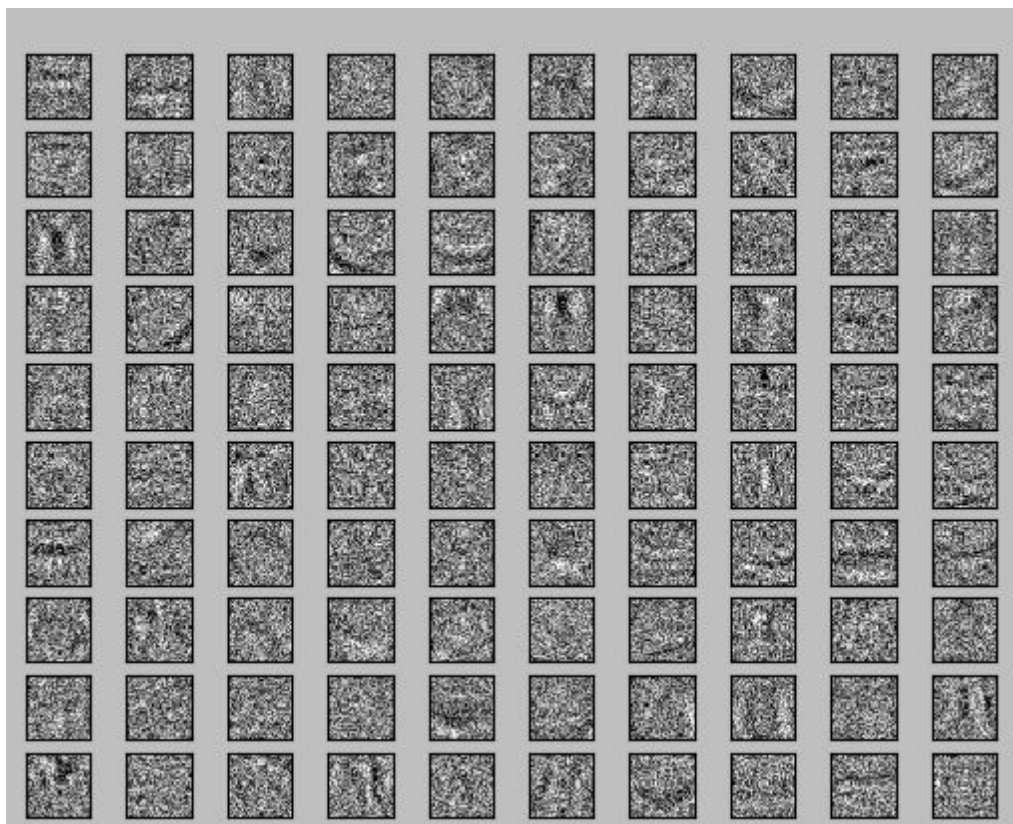
bez standaryzacji 0.278

Wynik dokładności w zależności do alpha.

Zmiany w alpie nie wniosły żadnych znaczących różnic



Wagi MLP w formie macierzy pikseli dla datasetu standaryzowanego



bez standaryzacji

