

Algorytm sieci neuronowej

Rozpoznawanie cyfr

Piotr Klimkowski Rafał Borysonek

Wydział Elektroniki
Politechnika Wrocławska

22 czerwca 2017

Spis treści

1. Teoria Sieci Neuronowych

1.1 Baza danych MNIST

1.2 Architektura sieci

1.3 Metoda Gradientu

1.4 Backpropagation

1.5 Overfitting

1.6 Dropout

2. Wyniki doświadczalne

2.1 Wpływ ilości ukrytych warstw na błąd

2.2 Wpływ ilości cykli uczenia na błąd

3. Prezentacja programu

4. Bibliografia

1. Teoria Sieci Neuronowych

1.1 Baza danych MNIST

1.2 Architektura sieci

1.3 Metoda Gradientu

1.4 Backpropagation

1.5 Overfitting

1.6 Dropout

2. Wyniki doświadczalne

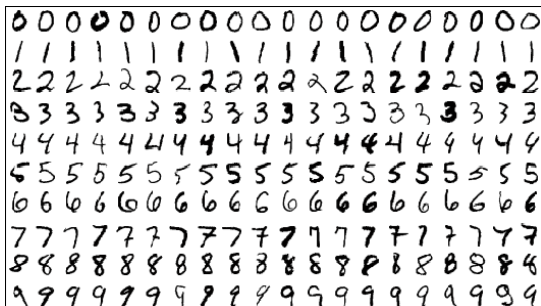
2.1 Wpływ ilości ukrytych warstw na błąd

2.2 Wpływ ilości cykli uczenia na błąd

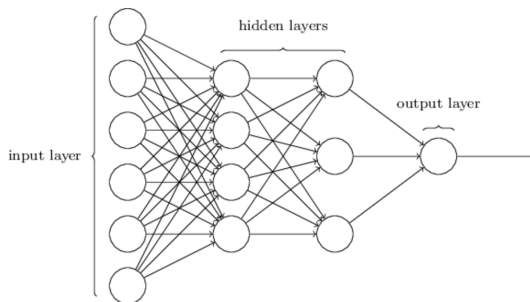
3. Prezentacja programu

4. Bibliografia

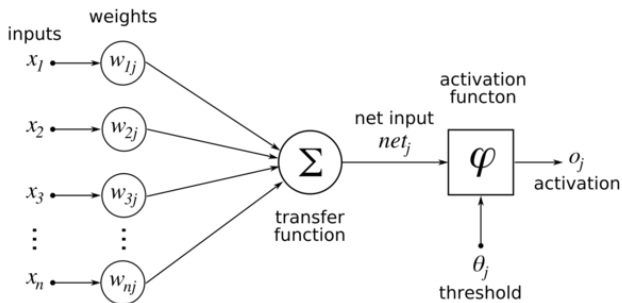
MNIST czyli baza danych służąca do uczenia:



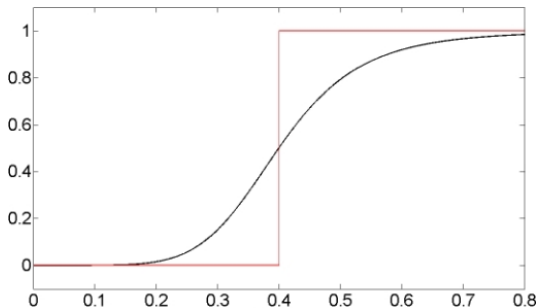
Struktura sieci



Neuron w sieci



Aktywacja neuronu

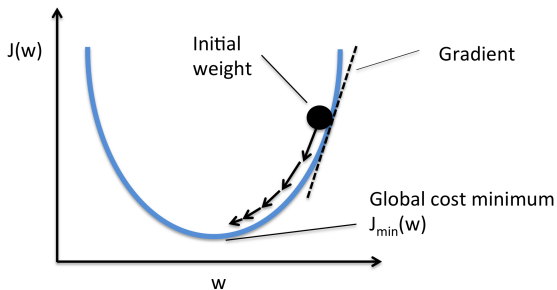


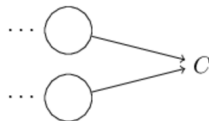
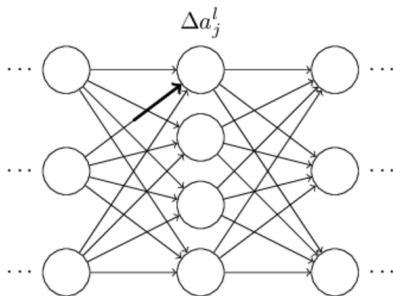
Dopasowywanie wag

Metoda najmniejszych kwadratów

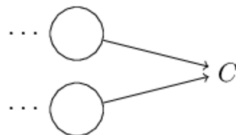
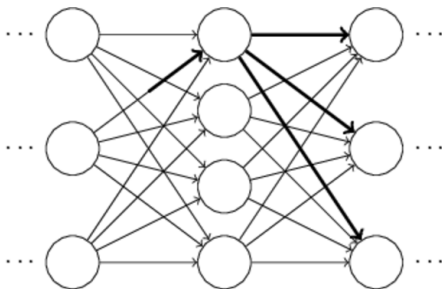
$$Q = \sum (y - \hat{y})^2$$

Gradient

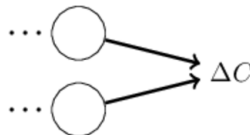
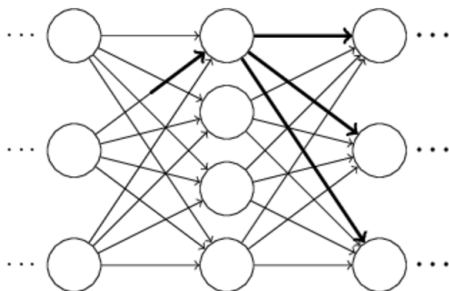




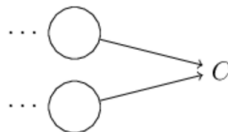
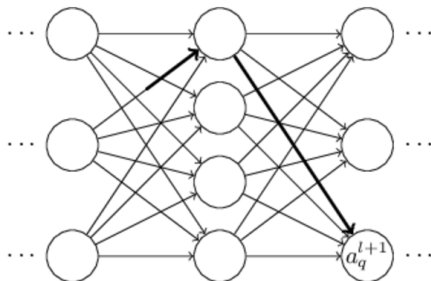
Backpropagation



Backpropagation



Backpropagation

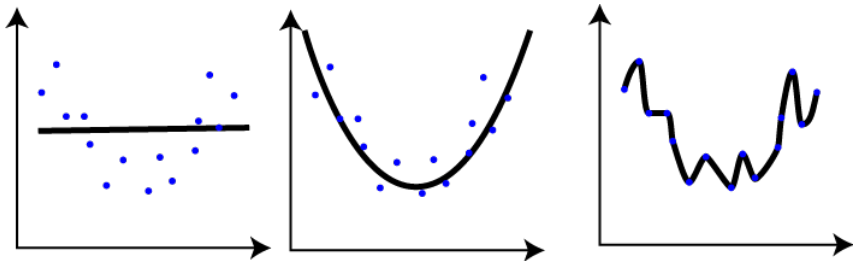


$$\Delta a_q^{l+1} \approx \frac{\partial a_q^{l+1}}{\partial a_j^l} \Delta a_j^l.$$

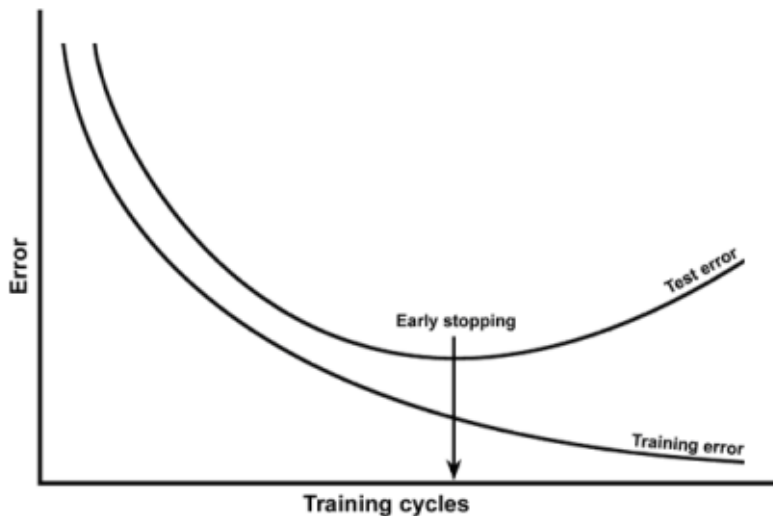
Backpropagation

$$\Delta C \approx \frac{\partial C}{\partial a_m^L} \frac{\partial a_m^L}{\partial a_n^{L-1}} \frac{\partial a_n^{L-1}}{\partial a_p^{L-2}} \cdots \frac{\partial a_q^{l+1}}{\partial a_j^l} \frac{\partial a_j^l}{\partial w_{jk}^l} \Delta w_{jk}^l,$$

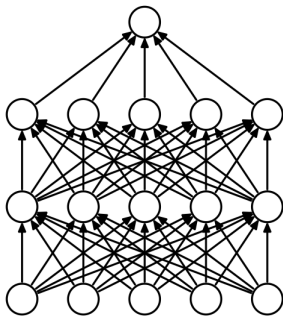
Co to jest Overfitting?



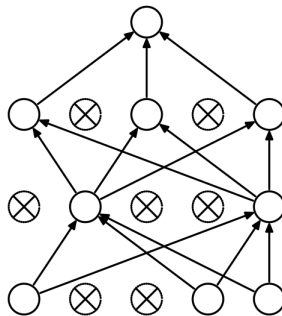
Kiedy on występuje?



Dropout jako sposób na overfitting



(a) Standard Neural Net



(b) After applying dropout.

1. Teoria Sieci Neuronowych

1.1 Baza danych MNIST

1.2 Architektura sieci

1.3 Metoda Gradientu

1.4 Backpropagation

1.5 Overfitting

1.6 Dropout

2. Wyniki doświadczałne

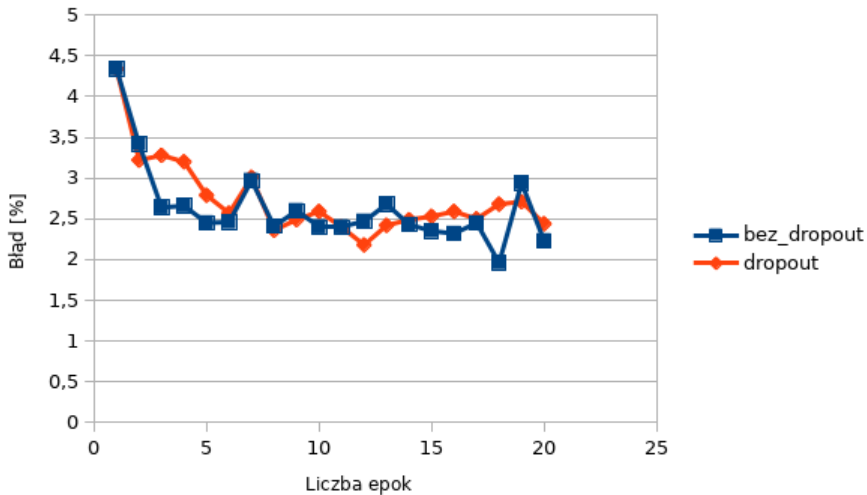
2.1 Wpływ ilości ukrytych warstw na błąd

2.2 Wpływ ilości cykli uczenia na błąd

3. Prezentacja programu

4. Bibliografia

Warstwy	Błąd [%]
800,400	2,32
400,800	2,40
800	2,41
400,400	2,42
500,150	2,52
800,800	2,53
400,200,100	2,53
800,400,400	2,61
400,400,400	2,62
800,400,100	2,94
100,100,100	3,23



1. Teoria Sieci Neuronowych

1.1 Baza danych MNIST

1.2 Architektura sieci

1.3 Metoda Gradientu

1.4 Backpropagation

1.5 Overfitting

1.6 Dropout

2. Wyniki doświadczalne

2.1 Wpływ ilości ukrytych warstw na błąd

2.2 Wpływ ilości cykli uczenia na błąd

3. Prezentacja programu

4. Bibliografia

1. Teoria Sieci Neuronowych

1.1 Baza danych MNIST

1.2 Architektura sieci

1.3 Metoda Gradientu

1.4 Backpropagation

1.5 Overfitting

1.6 Dropout

2. Wyniki doświadczalne

2.1 Wpływ ilości ukrytych warstw na błąd

2.2 Wpływ ilości cykli uczenia na błąd

3. Prezentacja programu

4. Bibliografia

Bibliografia I



Yann LeCun, Corinna Cortes, Christopher J.C. Burges.

THE MNIST DATABASE of handwritten digits.

<http://yann.lecun.com/exdb/mnist/>



Stephen Welch.

Neural-Networks-Demystified.

<https://github.com/stephencwelch>



Keras Documentation

Keras: The Python Deep Learning library.

<https://keras.io/>