

Deep Learning

Sieci typu Autencoder i praktyczne demo

Stanisław Jastrzębski

Kwiecien 2014

Plan prezentacji

- ▶ Problem klasyfikacji
- ▶ Sieci neuronowe
- ▶ Przykład w Pythonie

Definition

Machine learning to gałąź sztucznej inteligencji zajmująca się konstrukcją i badaniem systemów, które uczą się na podstawie danych.

Definition

Machine learning to gałąź sztucznej inteligencji zajmująca się konstrukcją i badaniem systemów, które uczą się na podstawie danych.

Deep Learning to jedna z gałęzi Machine Learningu

Rozpoznawanie obrazów

Klasyfikacja obiektów na obrazie to bardzo ważny problem

Przykłady zastosowań:

1. Rozpoznawanie pisma
2. Detekcja asteroidów
3. Medycyna
4. Wyszukiwanie grafiki (Google Images)
5. ...



Motocykl

Czemu to jest takie trudne

Co widzi człowiek, a kamera

Dla rozdzielczości 100x100 pixeli mamy 255^{10^4} możliwych obrazków.

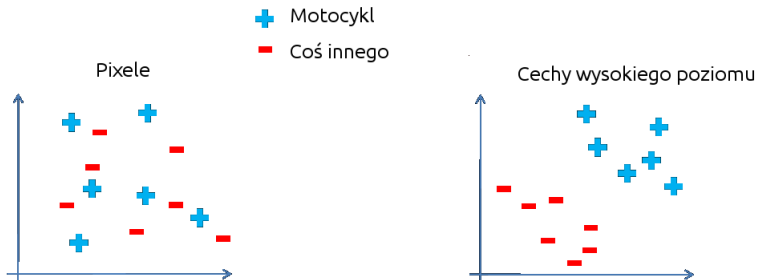


194	210	201	212	199	213	215	195	178	158	182	209
180	189	190	221	209	205	191	167	147	115	129	163
114	126	140	188	176	165	152	140	170	106	78	88
87	103	115	154	143	142	149	153	173	101	57	57
102	112	106	131	122	138	152	147	128	84	58	66
94	95	79	104	105	124	129	113	107	87	69	67
68	71	69	98	89	92	98	95	89	88	76	67
41	56	68	99	63	45	60	82	58	76	75	65
20	43	69	75	56	41	51	73	55	70	63	44
50	50	57	69	75	75	73	74	53	68	59	37
72	59	53	66	84	92	84	74	57	72	63	42
67	61	58	65	75	78	76	73	59	75	69	50

Pixełe nie wystarczają

Chcemy jak najlepiej rozdzielić dane

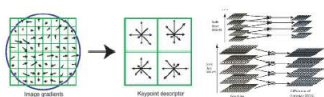
Mając dobrą reprezentację danych (na poziomie obiektów, a nie pojedynczych pixeli) wystarczy nauczyć model liniowy (np. regresję logistyczną - odmiana regresji liniowej).



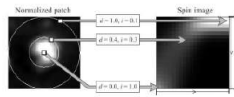
Reprezentacje danych

Jednym z możliwych rozwiązań jest stworzenie cech ręcznie. Są to detektory krawędzi, tekstur itp. Niestety:

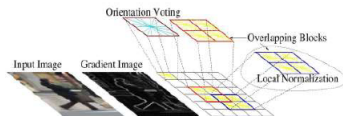
1. Dla każdego typu danych muszą być tworzone zupełnie od nowa
2. Nie są dostosowane do danego zestawu danych
3. Trudne w praktyce



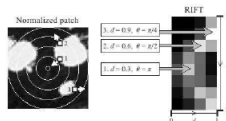
SIFT



Spin Image

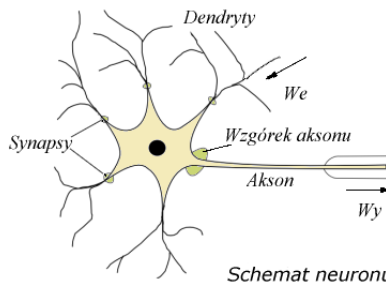
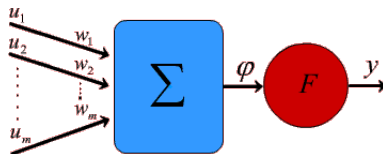


HoG



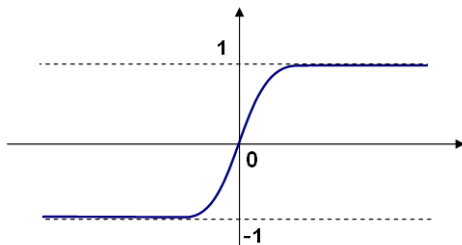
RIFT

Model neuronu



Model neuronu - aktywacja

$$f(x) = \sigma(< w, x >) \quad (1)$$



Autoenkoder

Połączmy neurony w sieć

Autoenkoder to sieć neuronów, która ma za zadanie **optymalnie** zakodować sygnał.

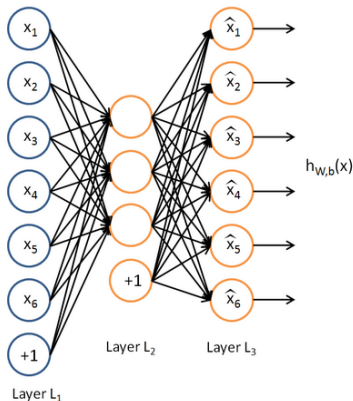
Autoenkoder

Połączmy neurony w sieć

Autoenkoder to sieć neuronów, która ma za zadanie **optymalnie** zakodować sygnał.

Aktywacja warstwy 2:

$$f(X^1) = \sigma(W^1 X^1) \quad (2)$$



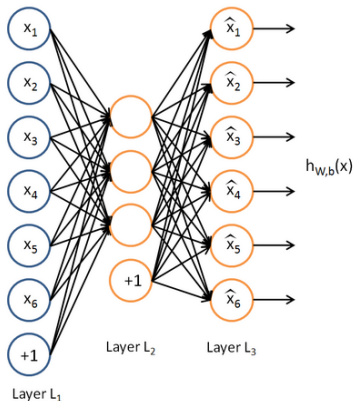
Autoenkoder

Połączmy neurony w sieć

Autoenkoder to sieć neuronów, która ma za zadanie **optymalnie** zakodować sygnał.

Aktywacja warstwy 2:

$$f(X^1) = \sigma(W^1 X^1) \quad (2)$$



Autoenkoder

Co to znaczy optymalnie?

$$J(W, b; x, y) = \frac{1}{2} \|h_{W,b}(x) - y\|^2. \quad (3)$$

Autoenkoder

Co to znaczy optymalnie?

$$J(W, b; x, y) = \frac{1}{2} \|h_{W,b}(x) - y\|^2. \quad (3)$$

Trickiem zapewniającym dobre działanie sieci jest dodawanie do wejścia (obrazka) losowego szumu.

Autoenkoder

Co to znaczy optymalnie?

$$J(W, b; x, y) = \frac{1}{2} \|h_{W,b}(x) - y\|^2. \quad (3)$$

Trickiem zapewniającym dobre działanie sieci jest dodawanie do wejścia (obrazka) losowego szumu.

Minimalizujemy koszt, czyli wybieramy takie wagi, że:

$$W^* = \operatorname{argmin}_{W,b}(J(W, b)) \quad (4)$$

Autoenkoder

Co to znaczy optymalnie?

$$J(W, b; x, y) = \frac{1}{2} \|h_{W,b}(x) - y\|^2. \quad (3)$$

Trickiem zapewniającym dobre działanie sieci jest dodawanie do wejścia (obrazka) losowego szumu.

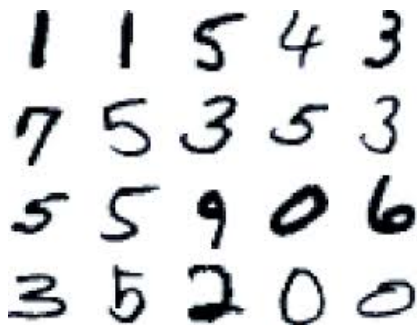
Minimalizujemy koszt, czyli wybieramy takie wagi, że:

$$W^* = \operatorname{argmin}_{W,b}(J(W, b)) \quad (4)$$

Minimalizacja zwykle przez metody gradientowe.

MNIST

Standardowy zbiór pisanych cyfr



Biblioteka Theano:

1. Automatyczne różniczkowanie
2. Kompilacja kodu do C
3. Możliwość uruchomienia na GPU
4. Optymalizacje numeryczne

Theano logo, featuring the word "theano" in a lowercase, blue, sans-serif font.

1. Słabo radzi sobie z przesunięciami i zmianami skali - brak standaryzacji i model nieodporny na zniekształcenia

1. Słabo radzi sobie z przesunięciami i zmianami skali - brak standaryzacji i model nieodporny na zniekształcenia
2. Rozpoznaje cyfry podobne do tych jakie widział - mały zbiór danych

1. Słabo radzi sobie z przesunięciami i zmianami skali - brak standaryzacji i model nieodporny na zniekształcenia
2. Rozpoznaje cyfry podobne do tych jakie widział - mały zbiór danych

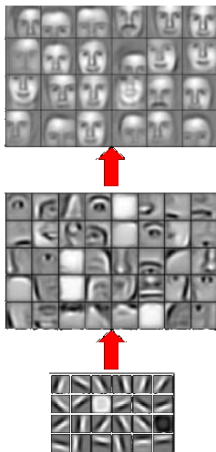
Powód: to jest bardzo prosty model (płytki i nauczony w 5 minut na małym komputerze).

1. Słabo radzi sobie z przesunięciami i zmianami skali - brak standaryzacji i model nieodporny na zniekształcenia
2. Rozpoznaje cyfry podobne do tych jakie widział - mały zbiór danych

Powód: to jest bardzo prosty model (płytki i nauczony w 5 minut na małym komputerze).

Prawdziwy model

Cechy wykryte przez głęboki model Google:



Dziękuję za uwagę