

Zastosowanie optymalizacji rojem cząstek (PSO) w procesie uczenia wielowarstwowej sieci neuronowej w problemie lokalizacyjnym

Jan Karwowski

Wydział Matematyki i Nauk Informacyjnych PW

17 XII 2013





Jan Karwowski, Michał Okulewicz, Jarosław Legierski (Orange Labs).
Application of particle swarm optimization algorithm to neural
network training process in the localization of the mobile terminal.
Engineering Applications of Neural Networks, strony 122–131.
Springer, 2013.



Problem

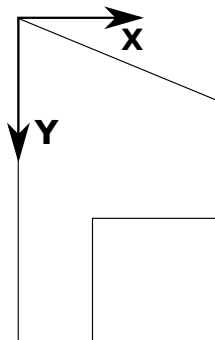
Wejście

- Wektor sił sygnałów z punktów dostępowych WiFi, 107 punktów dostępowych w budynku.
- Wektor sił sygnałów BTS sieci komórkowych

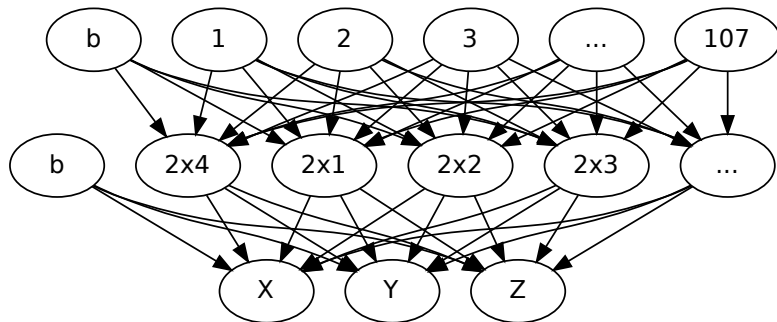
Około 1200 punktów pomiarowych, 40 pomiarów w punkcie w serii. 3 serie.

Wyjście

- Współrzędne XYZ
- Numer pomieszczenia
- Piętro



MLP



BP

Propagacja wsteczna błędów, aktualizacja wag po każdej prezentacji wzorca, w każdej iteracji zmiana kolejności wzorców.



M Clerc, i in.

Standard PSO 2011.

<http://www.particleswarm.info/>, Particle Swarm Central, 2011.

+

Restart co 50 iteracji wokół najlepszej znanej pozycji.

Mnożnik prędkości najlepszej cząstki






W każdym eksperymencie: Wejście: w wektor sygnałów jednego typu (WiFi lub GSM) przeskalowany do $[0,1]$.

Wyjście jeden z parametrów:

- ① Jako pojedyncza wartość przeskalowana do $[0,1]$ X , Y lub Z
- ② Jako wektor długości N , gdzie N , liczba klas — piętro lub pomieszczenie

Uśrednienie 40 obserwacji w punkcie.



-  Yuan-wei Jing, Tao Ren, Yu-cheng Zhou.
Neural network training using pso algorithm in atm traffic control.
Intelligent Control and Automation, strony 341–350. Springer, 2006.
-  Michael Meissner, Michael Schmuker, Gisbert Schneider.
Optimized particle swarm optimization (opso) and its application to artificial neural network training.
BMC bioinformatics, 7(1):125, 2006.
-  Xiaorong Pu, Zhongjie Fang, Yongguo Liu.
Multilayer perceptron networks training using particle swarm optimization with minimum velocity constraints.
Advances in Neural Networks–ISNN 2007, strony 237–245. Springer, 2007.



$$\mathbb{W} = [w_{0,0,1}, w_{0,0,2}, \dots, w_{0,n_0,n_1}, \dots, w_{l-1,n_{l-1},n_l}]$$

\mathbb{W} — wektor wszystkich wag w sieci l -warstwowej, n_i — liczba neuronów w warstwie i — tej, n_0 — liczba wejść, n_l — liczba wyjść, $w_{i,j,k}$ — waga połączenia między j — tym neuronem w warstwie i i k — tym w warstwie $i + 1$.

$$E : \mathbb{W} \rightarrow \mathbb{R}, E(W) = \sum_{p \in T} \frac{e(W, p)}{|T|}$$

$e(W, p)$ — $\frac{1}{2}$ MSE dopasowania wzorca p przez sieć o wagach W .
Optymalizujemy $E(W)$.

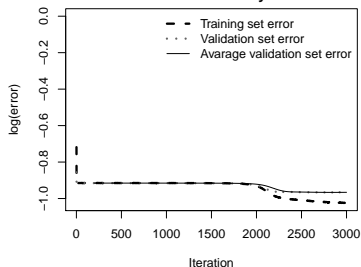


Zastosowanie PSO

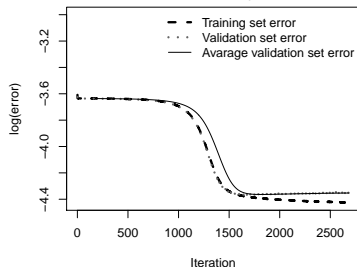
```
1: function PSOLEARNING(Network, Learning, Validation)
2:   Zainicjuj wszystkie cząstki losowo wokół środka układu
3:   for  $i = 1 \dots n$  do
4:     Wykonaj iterację PSO
5:      $best \leftarrow$  najlepsza znana pozycja od początku PSO
6:     if not VALIDATIONCHECK(Validation, best) then
7:       SETWEIGHTS(Network, best)
8:       return
9:     end if
10:  end for
11:  SETWEIGHTS(Network, best)
12:  BACKPROPAGATION(Network, Learning, Validation) ▷ Propagacja
    wsteczna z zatrzymaniem, gdy wzrośnie błąd walidacyjny
13: end function
```



Task: floor Hidden layers: 2

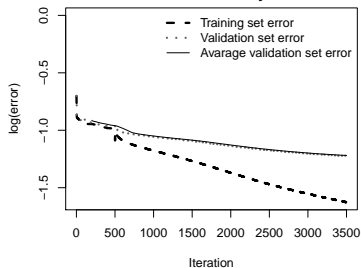


Task: x Hidden layers: 2

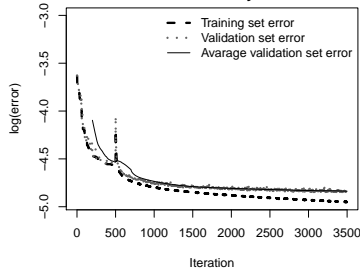


BP

Task: floor Hidden layers: 2



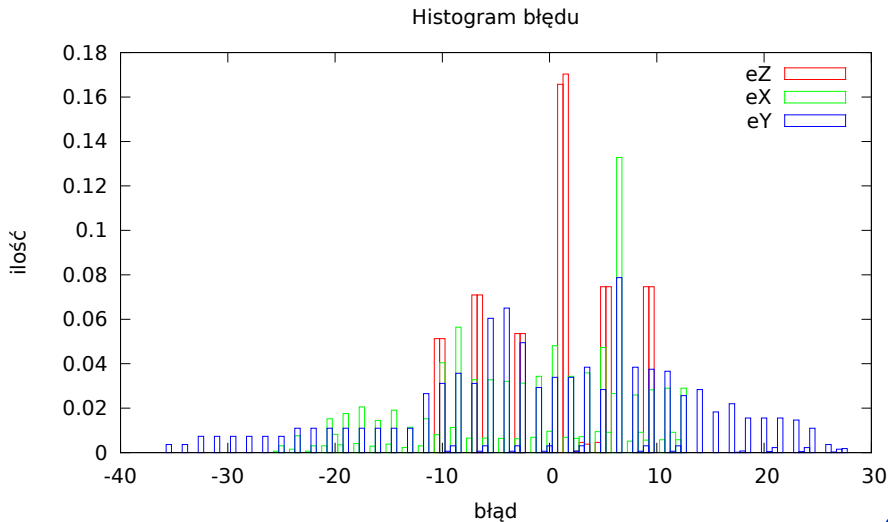
Task: x Hidden layers: 2



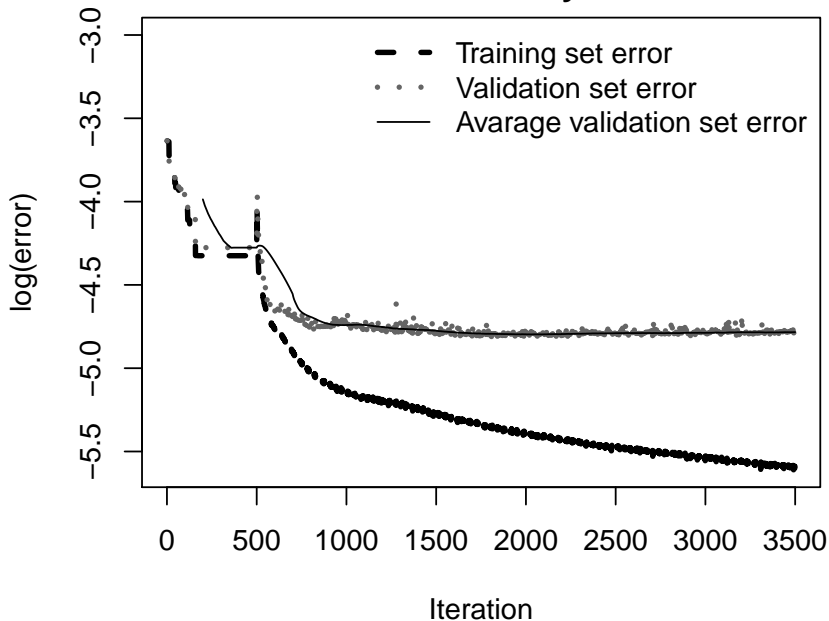
PSO



Początkowa faza uczenia BP



Task: x Hidden layers: 5



Skuteczność

PSO it.	Klasyfikacja		0.9 quantile of $ e_x $		0.9 quantile of $ e_y $	
	GSM	WiFi	GSM	WiFi	GSM	WiFi
5	54%	98%	8.66m	5.93m	11.98m	6.55m
10	53%	98%	8.91m	5.31m	11.08m	6.03m
25	53%	98%	8.37m	5.35m	11.40m	6.29m
50	53%	98%	8.80m	4.96m	12.87m	6.14m
100	52%	98%	8.61m	5.33m	11.76m	6.94m
150	56%	97%	9.13m	5.45m	12.47m	6.47m
200	54%	97%	8.90m	5.01m	11.92m	6.48m
250	53%	98%	8.77m	5.46m	11.55m	6.42m
500	54%	98%	9.11m	5.18m	11.66m	6.48m



Warstwy ukryte	Klasyfikacja		0.9 quantile of $ ex $	
	GSM	WiFi	GSM	WiFi
2	52%	98%	9.06m	5.54m
3	54%	98%	9.11m	5.18m
4	52%	98%	8.67m	5.37m
5	53%	98%	8.92m	5.27m



- PSO nadaje się do uczenia MLP
- Dobry punkt startowy dla BP
- Możliwość użycia sieci „z zapasem pojemności”
- Możliwość przetwarzania równoległego



- Inne grupowanie wzorców
- Zastosowanie innej miary błędu
- Nauka innych problemów
- Moment zmiany PSO→BP
- Usuwanie połączeń z sieci/ zmiany architektury w trakcie
- Modyfikacje PSO
- Powód dla którego błąd w BP jest początkowo stały



3 klasy

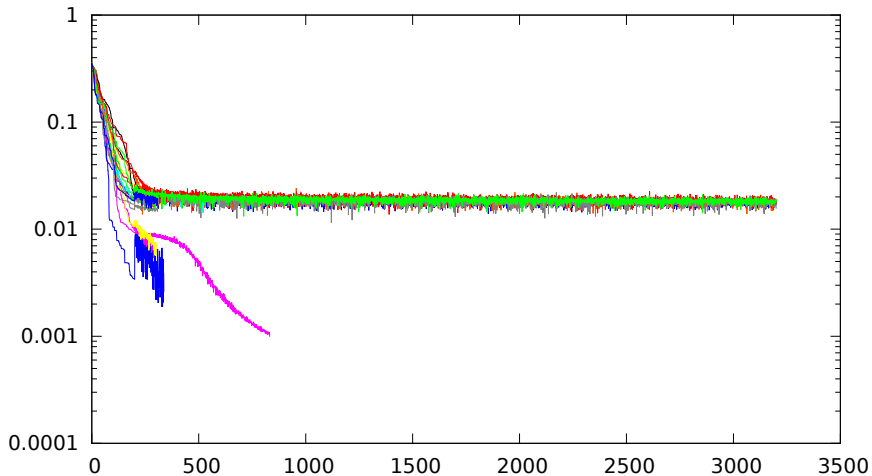
3 neurony w warstwie ukrytej

5 wejść

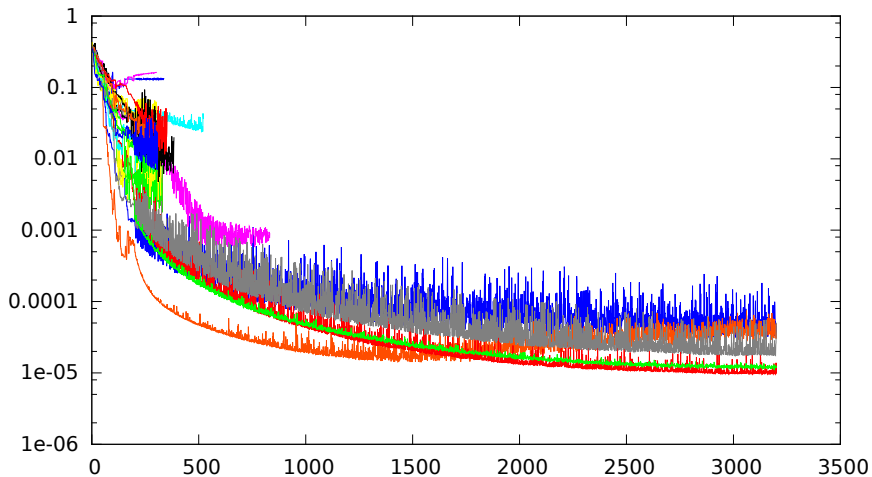
<http://archive.ics.uci.edu/ml/>

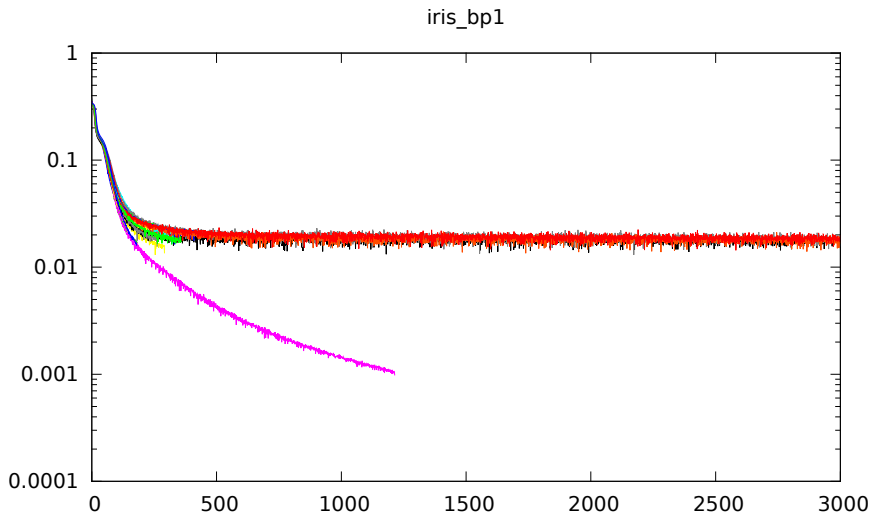


iris_pso1.iris_bp1



iris_pso1.iris_bp1 v





iris_bp1 v

