

Fakultet elektrotehnike i računarstva

ZAVRŠNI RAD br. 4699
Neuronske mreže izgrađene maxout-neuronima

Luka Žmegač

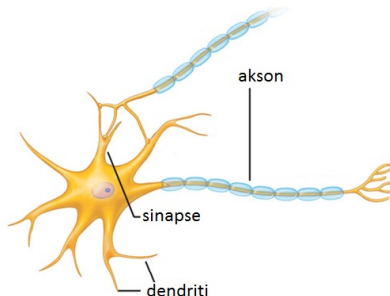
Zagreb, srpanj 2016.

Sadržaj

- 1 Umjetne neuronske mreže
 - Model neurona
 - Neuronske mreže tipa Maxout
- 2 Nadzirano učenje neuronskih mreža
 - Algoritam propagacije pogreške unatrag
 - Tehnika Dropout
- 3 Eksperimenti
 - Regresijski problemi
 - Klasifikacijski problemi

Neuroni

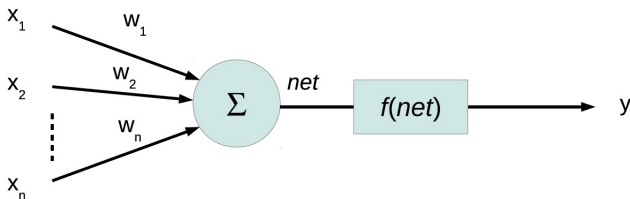
- ljudski mozak: 10^{11} neurona, 100 vrsta, 10000 veza
- spor rad neurona (vrijeme paljenja 1ms)
- → paralelan rad



Umjetni neuroni

- Definiran je:

- ulazima x_1 do x_n
- sinapse - težine w_1 do w_n
- tijelo neurona - funkcija ukupne pobude net
- akson - prijenosna funkcija $f(net)$



- Ukupna pobuda *net*:

$$net = w_0 \cdot x_0 + w_1 \cdot x_1 + \dots + w_n \cdot x_n$$

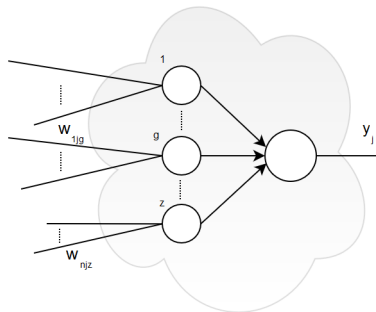
- Prijenosna funkcija modelira ponašanje i ulogu aksona.
 - Različite vrste: jedinična, funkcija skoka, linearna, po dijelovima linearna, sigmoidalna, ReLU...

Maxout-neuron

- Osmišljen kao prirodan par tehnici Dropout.
- Karakterizira ga prijenosna funkcija maxout.
- Sastoji se od grupe z linearnih dijelova. Grupe predstavljaju skupovi težina za razliku od običnih neurona koji imaju samo jedan skup težina.

Maxout-neuron

- Struktura Maxout-neurona vizualno prikazana u obliku skupine više neurona.
- Skupovi težina su prikazani kao zasebni neuroni.



Prijenosna funkcija maxout

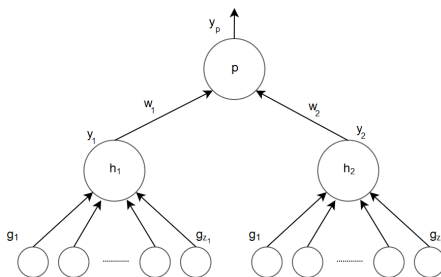
- Slična prijenosnoj funkciji ReLU
- Najveća vrijednost z dijelova odabire se kao aktivacija neurona. Prijenosnu funkciju neurona koji ima z skupova s n ulaza prikazujemo:

$$y = \max_{g \in 1 \dots z} net_g$$

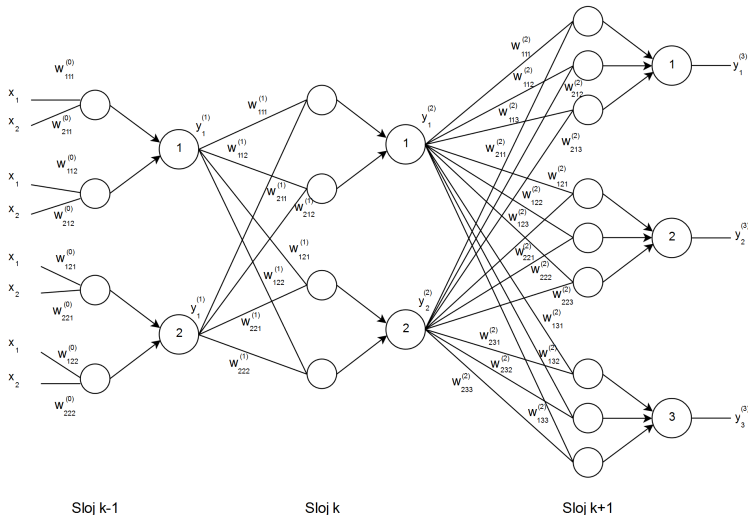
$$net_g = \sum_{i=0}^n w_{ig} \cdot x_i$$

Maxout-neuron kao aproksimator

- Struktura Maxout-neurona omogućuje mu da aproksimira bilo koju kontinuiranu funkciju koja je konveksna.
- Samo dva Maxout neurona mogu proizvoljno precizno oponašati bilo koju kontinuiranu funkciju, ako imaju dovoljno linearnih dijelova.



Primjer neuronske mreže tipa Maxout



Učenje neuronskih mreža

- Neuronske mreže moraju se trenirati (učiti) da bi se postigle željene funkcionalnosti mreže.
- Metode učenja:
 - Nadzirano učenje
 - Nenadzirano učenje
 - Potporno učenje
- Učenje mijenjanjem težina između neurona.

Algoritam propagacije pogreške unatrag

- Temelji se na računanju pogreške na izlazu te na njenoj propagaciji unatrag po mreži uz pomoć tehnike gradijentnog spusta.
- Kriterijska funkcija koja mjeri kvalitetu neuronske mreže definirana je kao polovica kvadrata srednje pogreške dobivenog i željenog izlaza.

$$E = \frac{1}{2} \sum_{o=1}^m (t_o - y_o^{(k+1)})^2$$

Algoritam propagacije pogreške unatrag

- Gradijent ovisan o prijenosnoj funkciji i tipu neurona.
- Po uzoru na gradijentni spust potrebno je izračunati gradijent od E .
- Težinski faktori $w_{ijg}^{(k)}$ mijenjat će se u skladu s gradijentnim spustom:

$$w_{ijg}^{(k)} \leftarrow w_{ijg}^{(k)} - \psi \cdot \frac{\partial E}{\partial w_{ijg}^{(k)}}$$

Algoritam propagacije pogreške unatrag primjenjen na neuronskim mrežama tipa Maxout

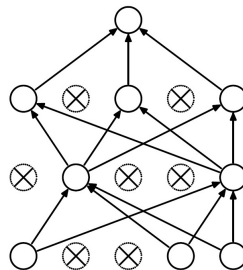
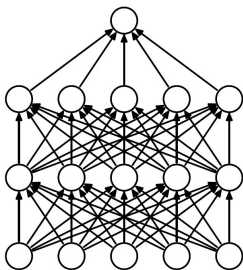
- Mreže tipa Maxout nemaju problem nestajućeg gradijenta.
- Potrebna vrlo mala stopa učenja.
- Različiti izrazi za promjene težina izlaznog sloja i skrivenih slojeva.
- Mijenjaju se samo težine skupa koji je generirao najveći *net* određenog neurona.

Tehnika Dropout

- Pokušava riješiti problem prenaučenosti (engl. *overfitting*) i neaktivnosti pojedinih neurona unutar neuronskih mreža.
- Moguće rješenje: učenje velikog broja različitih mreža.

Tehnika Dropout

- Pokušava riješiti problem prenaučenosti (engl. *overfitting*) i neaktivnosti pojedinih neurona unutar neuronskih mreža.
- Moguće rješenje: učenje velikog broja različitih mreža.
→ nije primjenjivo



- Privid učenja 2^n različitih mreža dijeljenih težina.
- Učenje - neuroni ulaznog i skrivenih slojeva ima vjerojatnost p za ispadanje iz mreže.
- Testiranje - izlaz neurona skalira se vrijednošću p .

Algoritam propagacije pogreške unatrag uz primjenu tehnike Dropout

- Koristi se algoritam propagacije pogreške unatrag s mini-grupama.
- Na početku računanja svake mini-grupe generiraju se izmjene mreže.
- Koristi se binarna maska.

Jednostavna linearna funkcija

$$f = \max \left\{ \begin{array}{l} -2x - 1 \\ 1.5x - 5.2 \end{array} \right\}$$

- Jedan Maxout-neuron, 2 skupa težina.
- Skup za učenje - 100 elemenata
- Dobivene težine neurona:
 1. skup težina

$$w_0 = -5.1968$$

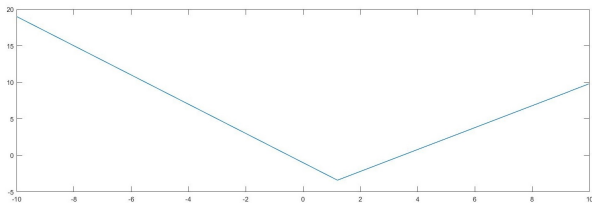
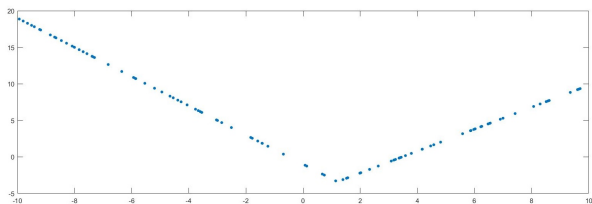
$$w_1 = 1.4997$$

2. skup težina

$$w_0 = -0.9999$$

$$w_1 = -1.9999$$

Regresijski problemi

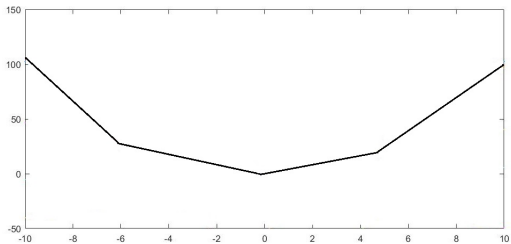
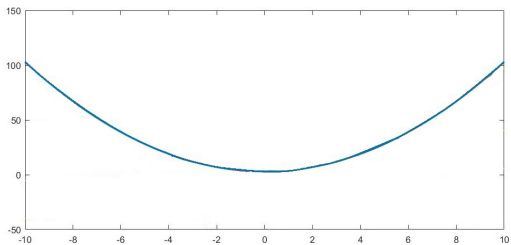


Parabola

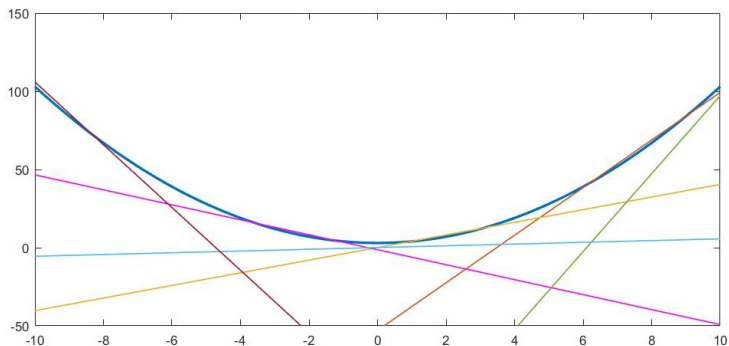
$$f = 1.2x^2 + 2.$$

- Jedan Maxout-neuron, 6 skupova težina.
- Skup za učenje: 100 elemenata.
- Aproksimacija konveksne funkcije pomoću linearnih djelova.

Parabola



Parabola

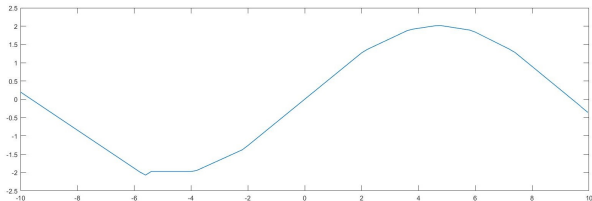
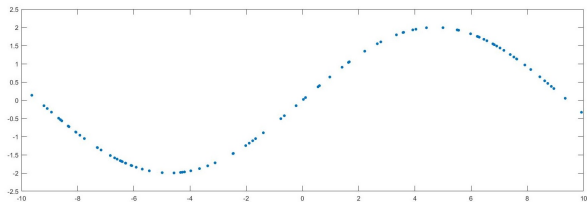


Sinusoidalna funkcija

$$f = 2 \cdot \sin(x/3).$$

- Dva sloja Maxout-neurona:
 - Ulazni sloj - 2 neurona, 15 skupova težina.
 - Izlazni sloj - 1 neuron, 2 skupa težina.
- Skup za učenje: 100 elemenata.
- Aproksimacija pomoću linearnih dijelova.

Sinusoidalna funkcija



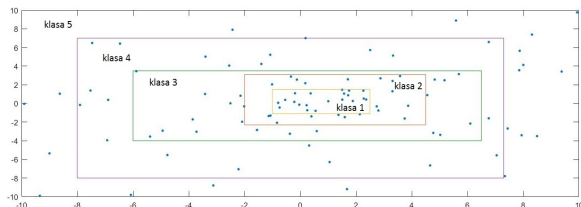
Klasifikacija

- Klasifikacija nije prirodan par mrežama tipa Maxout.
- Potrebno staviti sigmoidalnu funkciju na izlaz mreže.
 - Smanjenje iznosa gradijenta.

Klasifikacija ugniježđenog prostora

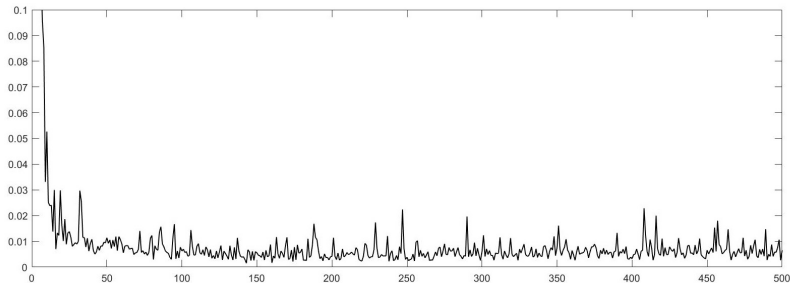
- Prostor u ovom primjeru je podijeljen u pet klasa koje su predstavljene ugniježđenim područjima.
- Tri sloja Maxout-neurona:
 - 1. sloj - 10 neurona, 6 skupova težina.
 - 2. sloj - 10 neurona, 5 skupova težina.
 - Izlazni sloj - 5 neurona, 3 skupa težina.
- Primjena tehnike Dropout.

Klasifikacija ugniježđenog prostora



- Srednja kvadratna pogreška $3,353 \cdot 10^{-6}$
- Omjer izlaznih vrijednosti koje nisu odstupale više od 5% od tražene vrijednosti (vrijednosti funkcije f) i traženih vrijednosti je 38/50.
- Točke u kojima mreža nije dobro klasificirala nalazile su se u graničnim područjima klasa koja mreža nije dobro odredila.

Srednja kvadratna pogreška



Zaključak

- Neuronske mreže tipa Maxout pokazale su se vrlo dobrima u problemima regresije, problemi pri klasifikaciji.
- Moguća poboljšanja: korištenje dodatnih tehnika pri učenju mreža, konvolucijske mreže.

Kraj

Hvala na pažnji

Izrazi

Izrazi za mijenjanje težina neuronske mreže algoritmom propagacije pogreške unatrag:

- težine izlaznog sloja:

$$w_{ijg}^{(k)} \leftarrow w_{ijg}^{(k)} + \eta \cdot \left(\sum_{s=1}^P \delta_{s,j}^{(k+1)} \cdot y_{s,i}^{(k)} \right)$$

$$\delta_{s,j}^{(k+1)} = (t_{s,j} - y_{s,j}^{(k+1)})$$

- težine skrivenih slojeva:

$$w_{ijg}^{(k)} \leftarrow w_{ijg}^{(k)} + \eta \cdot \left(\sum_{s=1}^P \delta_{s,j}^{(k)} \cdot y_{s,i}^{(k-1)} \right)$$

$$\delta_{s,j}^{(k)} = \sum_{o=1}^M \delta_{s,o}^{(k+1)} \cdot w_{jog_{max}}^{(k)}$$

Literatura I



Ian J. Goodfellow, Warde-Foley, D. M. Mirza, A. Courville, i Y. Bengio. Maxout networks. 2013.



Meng Cai, Yongzhe Shi, i Jia Liu. Deep maxout neural networks for speech recognition. 2013.



Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, i Ruslan Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. Journal of Machine Learning Research, 2014.



Marko Čupić, Bojana Dalbelo Bašić, i Marin Golub. Neizrazito, evolucijsko i neuroračunarstvo. 2013.