

Tema casa nr. 2

Retele neurale

Data afisarii: luni, 1 aprilie 2013

Termen predare: marti, 16 aprilie 2013, ora 23.59

Se accepta maxim 4 zile intarziere, cu penalizare (din 10 puncte) astfel:

- 0.5 puncte in prima zi;
- 1 punct /zi pentru urmatoarele 3 zile.

Tema se prezinta la laborator.

Limbaje de programare: C, C++, C#, Java, Scheme, Prolog, Lisp, Haskell, Python

Se considera o baza de date ce contine informatii despre persoane. Procentajul unei persoane poate fi determinat in functie de urmatoarele atribute:

- varsta (ani),
- greutate (lbs),
- inaltime (in)
- o serie de masuratori ale circumferintelor corpului (in).

Cerinte: [10 puncte]

Sa se implementeze o retea neurala de tip back propagation [1-4] care sa poata determina procentajul de grasime al unei persoane. Alegerea arhitecturii retelei (numar straturi ascunse, numar noduri/strat ascuns) va fi facuta prin *trial & error* [5].

Pentru antrenarea si testarea retelei se vor folosi datele din baza de date **BodyFat.csv**. In acest set de date, fiecare individ este descris prin masuratorile mentionate si are asociata clasa din care face parte.

Setul de date va fi impartit in :

- date de antrenare (2/3 din datele existente)
- date de test (restul datelor de test)

Antrenarea retelei se va face pe setul de date de antrenare. Se vor evalua performantele retelei, mai intai pe setul datelor de antrenare, iar apoi pe datele de test.

Pentru evaluarea performantelor retelei:

1. se va calcula eroarea absoluta si eroarea relativa pentru valorile din setul de antrenare si pentru valorile de test;
2. se vor afisa dependenta intre numarul de straturi ascunse / numarul de neuroni pe strat si:
 - eroarea absoluta/relativa pentru setul de date de test
 - numarul de pasi necesari pentru convergenta
3. se va afisa dependenta dintre numarul de pasi necesari pentru convergenta si viteza de invatare.
4. cerintele 2-3 vor fi rezentate grafic.

Observatie: Reteaua trebuie implementata. Nu se accepta folosirea simulatoarelor existente.

Bonus [2 puncte]: implementarea rețelei va fi făcută generic pentru a putea fi folosită și pe alte seturi de date. Pentru acest punct rețeaua va fi testată și pe setul de date **MPG.csv**.

References

[1]http://cortex.snowcron.com/neural_networks.htm

[2]<http://www.learnartificialneuralnetworks.com/backpropagation.html#understanding>

[3]<http://msdn.microsoft.com/en-us/magazine/jj658979.aspx>

[4]<http://ocw.mit.edu/courses/sloan-school-of-management/15-062-data-mining-spring-2003/lecture-notes/NeuralNet2002.pdf>

[5] <http://www.heatonresearch.com/articles/5/page2.html>