

NEURONSKE MREŽE: PREDVIĐANJE USPJEHA TELEVIZIJSKIH SERIJA KORIŠTENJEM NEUROPH STUDIO RAZVOJNOG OKRUŽENJA

ZADATAK: Korištenjem razvojnog okruženja Neuroph Studio potrebno je implementirati aplikaciju kojom se predviđa uspješnost (popularnost ili zarada) televizijskih serija. Neuronska mreža mora biti u mogućnosti odrediti izlazni parametar - uspješnost (slaba, srednja, visoka) prema zadanim ulaznim parametrima vezanim za npr. glavne glumce (npr. njihova zarada u zadnjih x filmova, broj uspješnih filmova/serija koje su snimili, itd.), redatelje, producente, raspoloživi budžet i sl. Može se uspoređivati i npr. u uspješnost pojedinih sezona (u kojoj sezoni se predviđa „gašenje“ serije). Broj ulaznih parametara mora biti minimalno 7, a izlaznih 3.

Zadatak je potrebno riješiti prema sljedećim koracima:

1. S mrežnih stranica potrebno je preuzeti Neuroph Studio razvojno okruženje:

<http://neuroph.sourceforge.net/download.html>

2. Tijekom implementacije rješenja potrebno je koristiti primjere sličnih aplikacija sa sljedeće mrežne stranice:

http://neuroph.sourceforge.net/sample_projects.html

Primjeri za usporedbu su svi koji u opisu imaju „**example of a multivariate data type classification problem**“, npr.

<http://neuroph.sourceforge.net/tutorials/SportsPrediction/Premier%20League%20Prediction.html>

<http://neuroph.sourceforge.net/tutorials/wines1/WineClassificationUsingNeuralNetworks.html>

3. Najveći dio zadatka je prikupljanje i organiziranje podataka za kreiranje skupa podataka za učenje – većina podataka o serijama dostupna je na <http://www.imdb.com/>
Dopušteno je korištenje i gotovih data setova (ako postoje i ako ih imate) kao i drugih izvora podataka.
4. Prema primjerima iz drugog koraka, procedura za učenje neuronske mreže se mora sastojati od normaliziranja podataka, kreiranja Neuroph projekta, kreiranje skupa za učenje, kreiranje neuronske mreže, samog postupka učenja mreže i provjeravanje ispravnosti rada mreže.

5. Na kraju je potrebno dokumentirati rezultate na primjerima već postojećih novijih serija (koji nisu korišteni u treningu!), kao i na serijama koje će tek izaći (tj. za koje još uvijek nemamo rezultate).
6. Predložak za izvještaje (dokumentaciju) su primjeri iz drugog koraka.

NAPOMENE:

- 1) Svi detalji koji nisu navedeni mogu biti proizvoljno definirani.
- 2) Studenti koji će uzimati najveće skupove podataka za učenje (smisleno organizirane) ostvariti će maksimalni broj bodova za laboratorijsku vježbu.