

## Neuralne mreže (13S053NM)

### Prvi projektni zadatak

#### 1 [6] Rešavanje problema klasifikacije primenom neuralne mreže

Varijantom D je data datoteka sa podacima koje treba klasifikovati primenom feedforward neuralne mreže. Prve dve kolone sadrže obeležja, dok je u trećoj koloni označena pripadnost klasi.

Arhitekturu neuralne mreže i aktivacione funkcije treba sami da definišete. Tokom obučavanja potrebno je isključiti sve zaštite od preobučavanja (rano zaustavljanje, regularizacija ...).

Podesiti da je fiksna slučajna raspodela (**rng(200)**).

Na graficima označiti ose na adekvatan način, i po potrebi dodati legendu. U izveštaju je potrebno opisati postupak rada i komentarisati dobijene rezultate.

- Vizualizovati podatke po klasama.
- Podeliti podatke na trening i test skup. Zašto je bitna podela podataka na ova dva skupa?
- Kreirati tri neuralne mreže koje će imati različite arhitekture: prva koja ne može da isprati dinamiku podataka (underfitting), druga koja na optimalan način klasifikuje podatke, i treća koja dovodi do efekta preobučavanja (overfitting). Za svaku arhitekturu definisati iste parametre treniranja: broj epoha, maksimalna dozvoljena greška, minimalan dozvoljen gradijent ...
- Izvršiti treniranje svakog modela i prikazati:
  - konfuzionu matricu na trening skupu,
  - konfuzionu matricu na test skupu,
  - granicu odlučivanja obučene mreže.
- Na osnovu dobijenih rezultata prokomentarisati performanse svakog od kreiranih modela. Međusobno uporediti uspešnosti dobijenih modela.

## 2 [7] Traženje optimalnih hiperparametara metodom unakrsne validacije

Varijantom S je data datoteka sa podacima koje treba klasifikovati primenom feedforward neuralne mreže korišćenjem unakrsne validacije.

Arhitekturu neuralne mreže i aktivacione funkcije treba sami da definišete. Tokom obučavanja optimizacije kriterijumske funkcije se vrši pomoću metode zadate varijantom M.

Podesiti da je fiksna slučajna raspodela (**rng(200)**).

Na graficima označiti ose na adekvatan način, i po potrebi dodati legendu. U izveštaju je potrebno opisati postupak rada i komentarisati dobijene rezultate.

- Opisati problem koji se rešava. Prikazati i komentarisati raspodelu odbiraka po klasama.
- Odvojiti podatke za treniranje i testiranje neuralne mreže. Objasniti kako je izvršena podela podataka. Kako raspodela podataka po klasama utiče na način podele podataka?
- Za svaku od metoda optimizacije su data 4 hiperparametra od kojih je potrebno izabrati 3 koja će se koristiti tokom unakrsne validacije. Navesti koji su hiperparametri uzeti u obzir, kratko opisati njihov uticaj na performanse mreže i navesti vrednosti koje će se koristiti tokom unakrsne validacije.
- Implementirati metodu unakrsne validacije za pronalaženje najboljih hiperparametara modela. Objasniti izbor mere performanse prema kojoj se ova procedura izvršava. Kako raspodela odbiraka po klasama utiče na izbor mere performanse. Navesti vrednosti optimalnih hiperparametara.
- Za finalno istreniran model prikazati:
  - krivu performanse neuralne mreže u zavisnosti od broja epoha,
  - matricu konfuzije na trening skupu,
  - matricu konfuzije na test skupu,
  - preciznost i osetljivost za klasu od interesa.
- Na osnovu dobijenih rezultata prokomentarisati performanse obučene neuralne mreže.

## Varijante ulaznih skupova

Studenti sa brojevima indeksa  $B_1B_1B_1B_1/G_1G_1G_1G_1$  i  $B_2B_2B_2B_2/G_2G_2G_2G_2$  rade ovaj zadatak sa parametrima:

$$D = \text{mod}(B_1B_1B_1B_1 + B_2B_2B_2B_2, 3) + 1;$$

$$M = \text{mod}(B_1 + B_1 + B_1 + B_1 + B_2 + B_2 + B_2 + B_2, 3) + 1;$$

$$S = \text{mod}(B_1 + B_1 + B_1 + B_1 + B_2 + B_2 + B_2 + B_2, 5) + 1;$$

D	Set podataka
1	dataset1
2	dataset2
3	dataset3

S	Set podataka	Link ka datasetu
1	CTG	<a href="https://www.kaggle.com/akshat0007/fetalhr">www.kaggle.com/akshat0007/fetalhr</a>
2	Genres	<a href="https://www.kaggle.com/mrmorj/dataset-of-songs-in-spotify">www.kaggle.com/mrmorj/dataset-of-songs-in-spotify</a>
3	JobChanges	<a href="https://www.kaggle.com/arashnic/hr-analytics-job-change-of-data-scientists">www.kaggle.com/arashnic/hr-analytics-job-change-of-data-scientists</a>
4	Rice	<a href="https://www.kaggle.com/mssmartypants/rice-type-classification">www.kaggle.com/mssmartypants/rice-type-classification</a>
5	Weather	<a href="https://www.kaggle.com/jsphyg/weather-dataset-rattle-package">www.kaggle.com/jsphyg/weather-dataset-rattle-package</a>

M	Metoda optimizacije	Hiperparametri
1	<i>trainscg</i>	arhitektura mreže, koeficijent regularizacije, težine klasa, aktivacione funkcije
2	<i>traingda</i>	arhitektura mreže, koeficijent regularizacije, težine klasa, konstanta obučavanja
3	<i>trainrp</i>	arhitektura mreže, koeficijent regularizacije, težine klasa, konstanta obučavanja