

## 13M051MU, 2. domaći zadatak 2025/26

### Generalizovani linearni modeli i gradijentni spust

Napišite kod za obučavanje softmax klasifikatora na osnovu podataka iz datoteke `multiclass_data.csv`. Poslednja kolona u ovoj datoteci je oznaka klase. Ostale kolone sadrže vrednosti prediktora. Koristite samo originalne prediktore (bez polinomijalne ekspanzije).

Za numeričku maksimizaciju verodostojnosti koristite stohastički gradijentni spust sa “mini-šaržama”. Kada koristimo gradijentni spust, tipično više puta prolazimo kroz čitav obučavajući skup. Jedan prolazak kroz ceo obučavajući skup se zove epoha. Uobičajeno je da se posle svake epohe podaci iz obučavajućeg skupa “izmešaju”, odnosno da redosled primera i njihovo grupisanje u mini-šarže bude slučajno izmenjeno na početku svake epohe.

Eksperimentišite sa stopom učenja  $\alpha$  i veličinom mini-šarže  $m_{mb}$ , i probajte da odredite kombinaciju ovih hiper-parametara koja dalje najbolju kombinaciju konačne ostvarene tačnosti i brzine konvergencije. Opciono, možete eksperimentisati sa naprednim tehnikama (inercija, AdaGrad, RMSProp, Adam). Ilustrujte konvergenciju graficima koji prikazuju gubitak na obučavajućem skupu i tačnost na validacionom skupu. Konkretno, treba da napravite dva skupa (jedan za tačnost na validacionom skupu i jedan za gubitak na trening skupu), sa po 5 grafika:

- jedan sa “optimalnom” kombinacijom vrednosti  $\alpha^*$  i  $m_{mb}^*$ ,
- dva sa  $\alpha^*$  i prevelikim, odnosno premalim  $m_{mb}$ ,
- dva sa  $m_{mb}^*$  i prevelikim, odnosno premalim  $\alpha$ .

Na  $x$ -osama grafika treba da bude redni broj epohe, a na  $y$ -osama prosečni gubitak tokom epohe (ukupan gubitak na svim primerima u trenutnoj epohi, podeljen brojem primera u obučavajućem skupu), odnosno ostvarena tačnost na kraju epohe na validacionom skupu.

Dodatno, treba da izmerite prosečno trajanje epohe obučavanja za svaku kombinaciju  $(\alpha, m_{mb})$ . Koristite metodu `time()` iz istoimenog Python modula. Ova vremena prikazite tabelarno, tako da vrste odgovaraju stopama učenja, a kolone veličinama mini-šarže.

Izračunajte i ukupna vremena do konvergencije, kao proizvod prosečnog trajanja epohe i broja epoha do dostizanja konačne vrednosti (procenite sami kada je prvi put dostignuta vrednost koja je bliska konačnoj). Njih takođe prikazite u odgovarajućoj tabeli.

Nije dozvoljeno korišćenje gotovih f-ja za softmax niti gradijentni spust. Python kôd i kratak izveštaj sa traženim graficima i tačnostima u pdf formatu, predati putem Teams-a do datuma naznačenog u sekciji Assignments. Kôd i izveštaj možete objediniti u Jupyter Notebook; u tom slučaju pošaljite ipynb i njegovu html ili pdf verziju sa generisanim rezultatima. Na kraju izveštaja navedite ostvarene tačnosti na trening skupu za oba korišćena modela.