

# 13M051MU Mašinsko učenje

## 1. domaći zadatak 2023/24: Regresija

### Podaci

U datoteci `data.csv` su podaci za jedan regresioni problem. Format datoteke je sledeći:

- prvih 5 kolona su prediktori;
- poslednja kolona je ciljna promenljiva;
- svaka vrsta predstavlja jedan obučavajući primer.

### Zadatak

Model koji treba da implementirate određen je zbirom  $S$  cifara vašeg broja indeksa, po modulu 3:

$S \bmod 3$	Model
0	grebena ( <i>ridge</i> ) regresija
1	LASSO regresija
2	lokalno ponderisana linearna regresija

Na primer, za broj indeksa  $20^{**}/3221$  je

$$S = 3 + 2 + 2 + 1 = 8, \quad S \bmod 3 = 2$$

pa je metod lokalno ponderisana linearna regresija. Kod grebene i LASSO regresije za hipotezu usvojite polinom 2. reda. Optimalnu vrednost hiper-parametra odredite unakrsnom validacijom.

### Rešenje

Preporuka je da se domaći zadatak radi u Python-u, a toleriše se i Matlab/Octave. Rešenje podrazumeva sledeće elemente:

1) izveštaj (pdf ili html) sa

- grafikom koji demonstrira kako ste izabrali hiper-parametar vašeg modela,
- izabranom vrednošću hiper-parametra,
- procenom vrednosti korena srednje-kvadratne greške vašeg modela.

2) kompletan kôd.

Razmislite kako biste mogli implementirati “produkcijski” model, koji bi primao jednu ili više petorki prediktora, a vraćao odgovarajuće procenjene vrednosti ciljne promenljive.

### Napomene

Nije dozvoljeno korišćenje gotovih f-ja za linearnu, polinomijalnu, grebenu, LASSO i lokalno ponderisanu regresiju, niti za validaciju, osim u cilju provere ispravnosti sopstvenih rezultata. Sve f-je treba da napišete sami, koristeći samo konstrukte za rad sa vektorima i matricama, optimizaciju, generisanje slučajnih brojeva, crtanje grafika itd (Python moduli numpy, scipy, matplotlib, itd).

Kôd i izveštaj predajete putem Teams-a do naznačenog datuma. Upustvo za predaju zadatka putem MS Teams-a možete naći ovde.