

Studijski program:

Datum:

Broj indeksa:

Ime i prezime:

Prvi računarski kolokvijum iz predmeta

Numerički algoritmi i numerički softver

Grupa A_1

Dat je skup podataka o studentima. Cilj je napraviti model linearne regresije koji na osnovu podataka iz *data/train.csv* fajla predviđa uspeh studenata na testu (kolona "Test Score").

Osim ako nije drugačije naglašeno, za metriku u zadacima koristiti prilagođeni r^2 , a sve statističke testove sprovesti sa pouzdanošću od 95%.

Zadatak 1.

Napraviti model linearne regresije koji koristi sve atribute. Ispisati meru koju model ostvaruje.

Zadatak 2.

Proceniti koji je minimalni rezultat na testu (*Test Score*) studenta koji je spavao 8 sati. Procenu izvršiti nad novim modelom koji na osnovu kolone *Sleep Hours* predviđa *Test Score*.

Zadatak 3.

Unaprediti model tako da su sve pretpostavke zadovoljene i prilikom testiranja je ostvarena mera prilagođeni $r^2 \geq 0.42$. Koristiti train/val podelu u odnosu: 70/30. Ispisati meru koju model ostvaruje.

Napomena: koristiti `random_state = 123` (kako bi se rezultati mogli reprodukovati).

Zadatak 4.

Objasnite šta su reziduali modela.

Zadatak 5.

Objasnite zašto je regresija pomoću polinoma trećeg stepena takođe linearna regresija.