Poštovane koleginice i kolege,

Zahtevi za drugi domaći zadatak su:

1) Učitati skup podataka Bank Marketing (skup podataka i opis atributa se nalazi u prilogu)

2) Analizirati skup podataka

- Odrediti tipove podataka atributa,

- Raspodele vrednosti,

- Nedostajuće vrednosti,

- Deskriptivne statistike,

- Korelacije,

- Vizualizacije podataka, i slično.

3) Pripremiti skup podataka tako da bude pogodan za prediktivno modelovanje

- Popuniti nedostajuće vrednosti,

- Kvantifikovati kategoričke podatke,

- Kreirati izvedene atribute

**Klasifikacija:**

5) Naučiti najmanje tri prediktivna modela (sa uobičajnim postavkama parametara) i

- Oceniti kvalitet modela na trening skupu podataka,

- Oceniti kvalitet modela unakrsnom validacijom na trening skupu podataka.

- Prilikom ocenjivanja kvaliteta modela koristiti više mera evaluacije modela, i prodiskutovati cene grešaka, kao i matricu konfuzije.

- Diskutovati o pretreniranosti modela.

6) Prilagoditi parametre prediktivnih modela i uporediti kvalitet modela pre i posle promene parametara.

7) Prilagoditi granicu odlučivanja prediktivnih modela i uporediti kvalitet modela pre i posle promene granice.

**Klasterovanje:**

8) Identifikovati i karakterisati ekstremne vrednosti (*outlier-e*)

9) Kreirajte klaster model i odredite klastere svake instance

10) Karakteristati dobijene klastere

11) Ocenite kvalitet klaster modela

12) Odrediti najbolji broj klastera koristeći mere kvaliteta klaster modela

**BONUS 1**: Podeliti inicijalni skup podataka po klasterima koje ste dobili (ili u zadatku 9 ili u zadatku 12) i sačuvajte ih u posebnim promenjivim. Na svakom od tako dobijenih podskupova podataka naučite barem jedan prediktivni model i uporedite rezultate između podskupova podataka i sa rezultatima dobijenim pre deljenja skupa podataka po klasterima.