

Mašinsko učenje, Januar1 (PS) 2025. godine

Na Desktop-u možete pronaći arhivu sa imenom *ML_jan1_2025_materijali.zip* u kojoj se nalaze materijali potrebni za rad. Raspakovati arhivu pa dobijeni direktorijum preimenovati tako da odgovara vašim podacima u formi *ML_jan1_2025_ImePrezime_BrojIndeksa*. Zatim mu pristupiti iz terminala pokretanjem komande *jupyter notebook*.

Na Desktop-u se nalazi i direktorijum sa imenom *docs* u kojem se nalazi dokumentacija.

1. (25 poena) Regresija.

U datoteci *data/concrete_data.csv* nalaze se podaci o uzorcima betona. Jedno od najvažnijih svojstava betona koje se određuje eksperimentalno je njegova čvrstoća, odnosno otpornost betona na pritisak, koja se izražava u *MPa* (megapaskalima) i zavisi od sastava i starosti betona. Podatke iz datog skupa podataka treba iskoristiti za predikciju ove vrednosti. Kratak pregled skupa:

- *Cement* - cement
- *Blast Furnace Slag* - šljaka visoke peći
- *Fly Ash* - leteći pepeo
- *Water* - voda
- *Superplasticizer* - superplastifikator, aditiv
- *Coarse Aggregate* - krupni agregat, šljunak
- *Fine Aggregate* - fini agregat, pesak
- *Age* - starost izražena kao broj dana
- *Concrete compressive strength* - pritisna čvrstoća betona

Svi atributi osim starosti predstavljaju komponente u sastavu betona i njihove vrednosti su izražene u kg/m^3 .

Izvršiti analizu i preprocesiranje atributa, a zatim obučiti dva klasična modela mašinskog učenja (klasični modeli ne uključuju neuronske mreže) koji predviđaju čvrstoću betona.

Uporediti dva obučena modela u terminima regresionih metrika.

2. (25 poena) Klasifikacija. Duboko učenje.

U datotekama *carboard.npy*, *metal.npy*, *glass.npy* i *plastic.npy* se nalaze slike otpada za reciklažu. Napraviti klasifikator otpada korišćenjem konvolutivne neuronske mreže sa tačnošću od bar 55%. Za učitavanje podataka koristiti `np.load` funkciju. Učitani podaci će biti oblika (H, W, C), a kako `torch` očekuje broj kanala pre širine i visine, koristite funkciju `torch.transpose`. Iscrtati krive funkcije greške i tačnosti modela u toku obučavanja mreže i ispisati klasifikacione metrike najboljeg modela.

Rešenja zadataka sačuvati kao Jupyter sveske *1.ipynb* i *2.ipynb*. Sveske pokrenuti i sačuvati u direktorijumu *src*.