

## Istraživanje podataka - Vežbe 6

1. Na Desktopu se nalazi skup podataka *car.csv* o klasama automobila. Koristeći alat KNIME i neuronske mreže (MLP) izvršiti klasifikaciju nad datim skupom. Atribut *class* sadrži informaciju kojoj klasi pripada automobil. Napraviti različite modele klasifikacije sa različitim arhitekturama neuronske mreže promenom broja skrivenih slojeva i skrivenih čvorova. Broj skrivenih slojeva je u intervalu [1, 3], a broj skrivenih čvorova po sloju je u intervalu [5, 9]. U okviru radnog toka napraviti tabelu koja za svaki model sadrži sledeće podatke:

- preciznost na trening skupu;
- preciznost na test skupu;
- broj skrivenih slojeva;
- broj skrivenih čvorova po sloju;
- model u PMML formatu;

Odgovoriti na pitanja/ uraditi zadatke:

- Uporediti dobijene modele prema preciznosti.
- Kolika je preciznost na trening i test skupu za model sa najvećom preciznošću na test skupu?
- Opisati model sa najvećom preciznošću na test skupu.
- Za model sa najvećom preciznošću na test skupu nacrtati neuronsku mrežu.

2. Koristeći skup podataka *car.csv* izvršiti klasifikaciju metodom KNN primenom unakrsne validacije u programskom jeziku Python. Napraviti različite modele klasifikacije promenom parametara broj najbližih suseda, mere za računanje rastojanja između instanci i funkcije težine za susede. Izdvojiti izveštaj o klasifikaciji za najbolji model prema preciznosti i prikazati matricu konfuzije za trening i test skup. Opisati dobijeni model.