**Jan Lampič**

**1.Domača naloga pri predmetu ITAP**

Preden sem lahko začel s treniranjem modelov na učni množici podatkov, sem mogel dane podatke najprej urediti. Stolpcema X1 in X3 je bilo potrebno spremeniti tip v numeric, nato pa sem odstranil še vse ne numerične meritve.

Za treniranje modela je bilo potrebno definirati novo mero natančnosti .

.

Začel sem z najbolj osnovnim modelom (metode najbližjih sosedov) ***knnModel****.*Metodo sem pognal večkrat in vsakič z drugačnim parametrom *k* oz. drugačnim številom sosedov. Optimalno rešitev sem dosegel pri *k* = 7 in sicer = 0.64.

Za naslednji model sem uporabil logistično regresijo ***logModel***. Da bi se čimbolj približal pravemu modelu, sem logistični regresiji dodajal tudi člene višjih redov (vse do 5. stopnje), vendar se je to izkazalo za neučinkovito, saj je bila najbolj učinkovita logistična regresija s členi prvega reda ( = 0.68).

Za zadnji model sem uporabil odločitveno drevo ***rpartModel***. Tu sem uporabil knjižnico **rpart** za učenje odločitvenega drevesa, vendar tudi njegova natančnost ni bila boljša od linearnega modela ( = 0.64). Za napovedni model sem potem izbral logistično regresijo, saj je imela ta največjo vrednost.

Mislim, da bo vrednost na moji napovedi podobna tisti, ki sem jo dobil pri modelu logistične regresije oz. okoli 0.68, saj sem pri testiranju vrednost dobil s pomočjo prečnega preverjanja. Ocena bi bila lahko mogoče malo boljša, če bi pri testiranju uporabil vzorčenje (vendar je to računsko zahtevnejše).