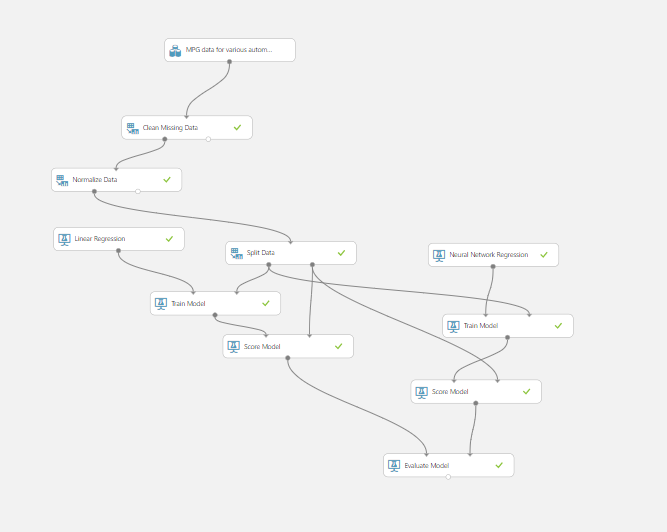
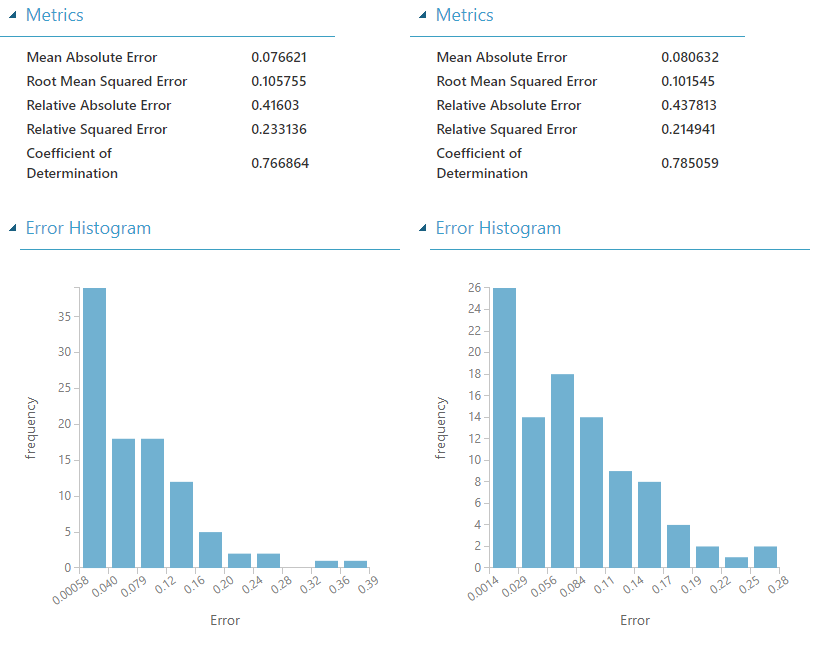
**RUAP – LV6: Alati i usluge za analizu podataka**

**David Kvesić - DRC**

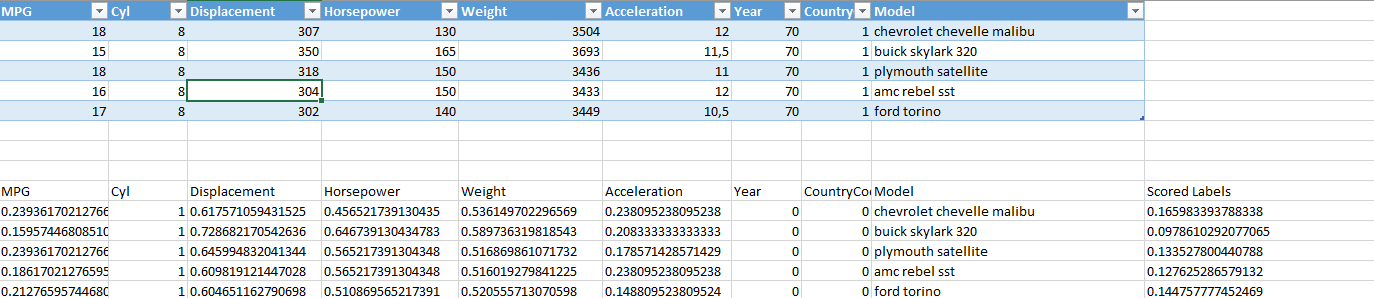
**Repozitorij :** [**https://github.com/Kvesss/RUAP-LV6**](https://github.com/Kvesss/RUAP-LV6)

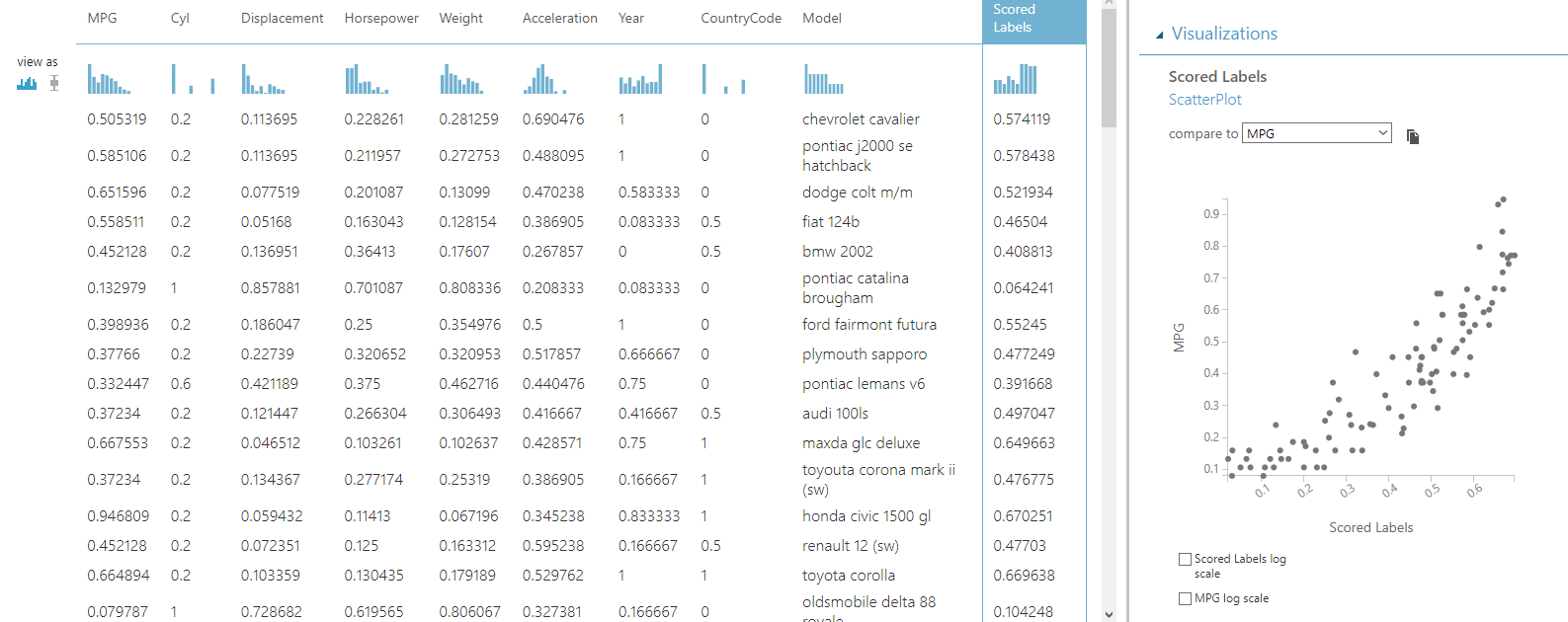
**Zadatak 1.**



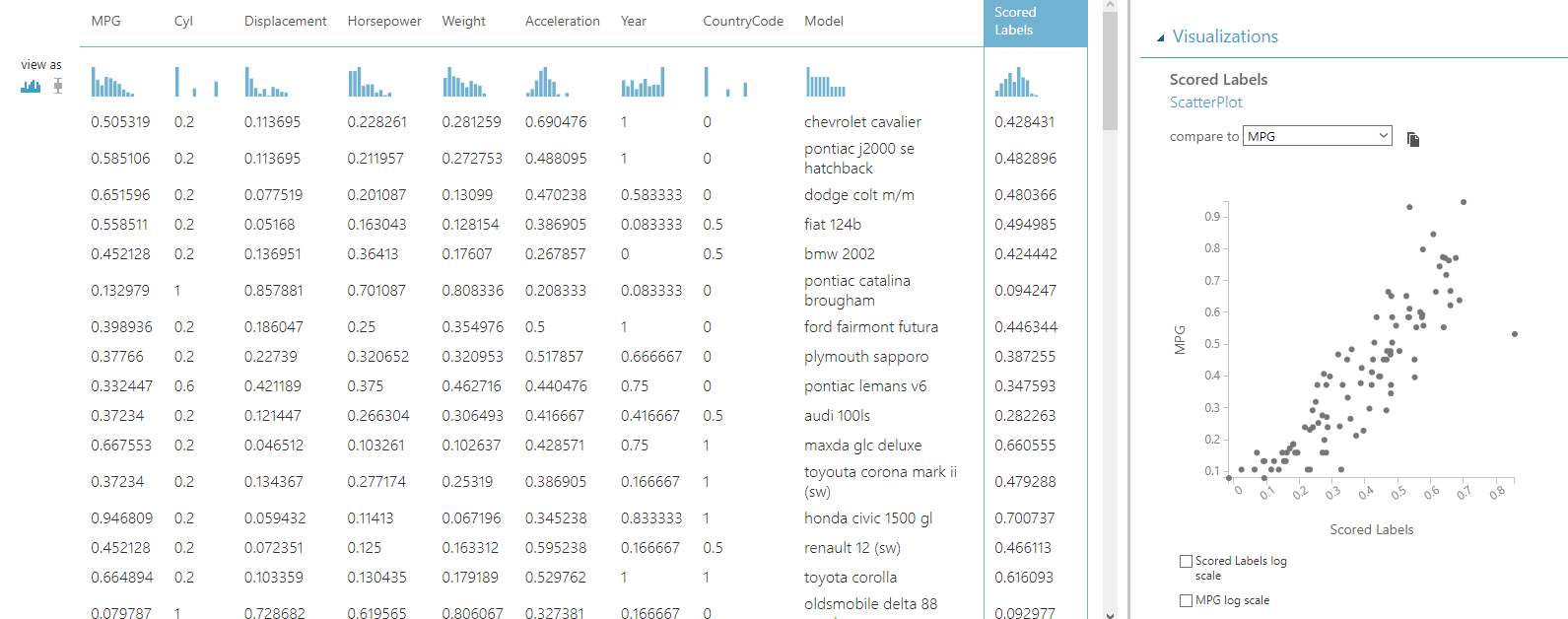


Linearna regresija (lijevo) i neuronska mreža (desno)





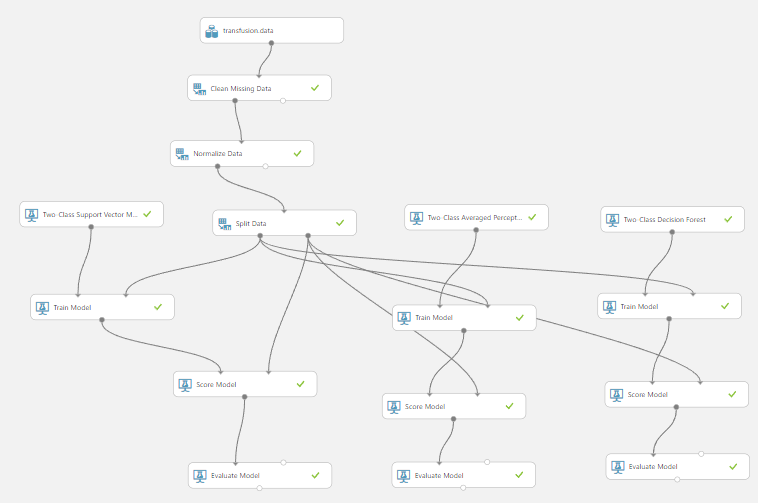
ANN

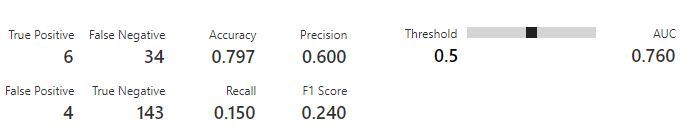


Linear Regression

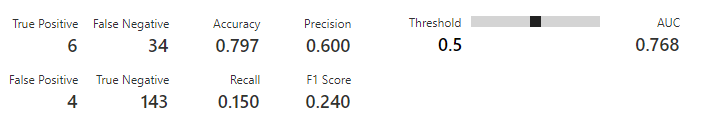
Metoda linearne regresije I metoda neuronskih mreža daju slične rezultate. Srednja apsolutna pogreška je manja kod linearne regresije, a koeficijent determinacije je nešto veći u slučaju kNN. Zbog sličnih rezultata u ovom slučaju je teško reći koji je model bolji.

**Zadatak 2.**

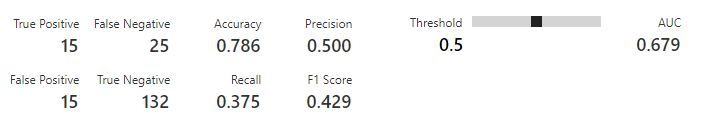




Two-Class Support Vector Machine



Two-Class Averaged Perceptron



Two-Class Decision Forest

Korišteni algoritmi klasifikacije su Two-Class Support Vector Machine, Two-Class Averaged Perceptron i Two-Class Decision Forest. Sa i bez normalizacije model je dao iste rezultate što znači da primjena normalizacije u ovoj situaciji nema značajnu ulogu.

Rezultati pokazuju da su prva dva algoritma dali identične rezultate za preciznost ( “Accuracy” ), dok je Two-Class Decision Forest bio neznatno manje precizan. Ipak F1 score je za taj algoritam veći od prva dva što znači da je taj algoritam u ovom slučaju najbolji jer je F1 score neovisan o razdiobi ulaznih podataka, a u ovom slučaju znatno je veći broj ljudi koji nisu dali krv sa vrijednosti “0” od onih koji su dali krv odnosno “1”. To se najbolje vidi na primjeru True positive rezultata gdje su prva dva algoritma točno predvidjela samo 6 od 40 negativnih slučajeva (15%) dok je treći imao 15 od 40 (37.5%). Iako su prva dva algoritma bolje predvidjela True negative slučajeve (97.3% u odnosu na 89.8%) ta je razlika puno manja u odnosu na prethodnu kod true positive rezultata.