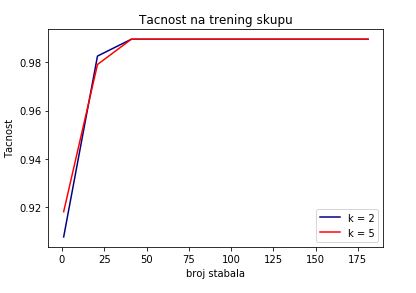
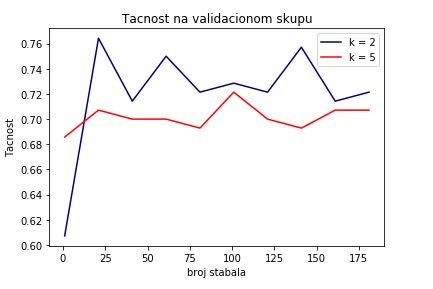
**ANSAMBLI**

1. Random Forest – ispitivanje maksimalnog broja odlika koje se razmatraju pri generisanju čvorova stabla u zavisnosti od veličine ansambla.

Ispitivane su vrednosti 2 i 5 I dobijeni su sledeći rezultati (kod u fajlu RF\_k.py):

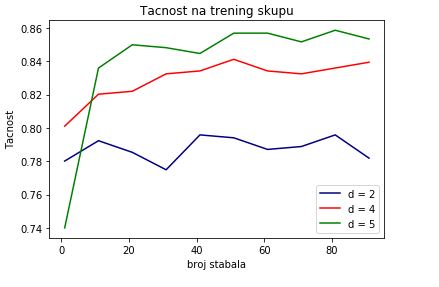


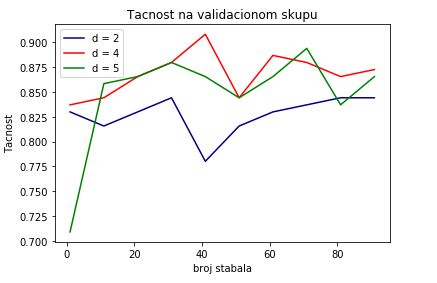


Zaključak: manji parameter *k* smanjuje međuzavisnosti stabala u ansamblu čime obezbeđuje bolji rezultat pri ansambliranju.

1. Random Forest – ispitivanje maksimalne dubine stabla u zavisnosti od veličine ansambla.

Ispitivane su dubine 2, 4 i 5 I dobijeni su sledeći rezultati (kod u fajlu RF\_depth.py):

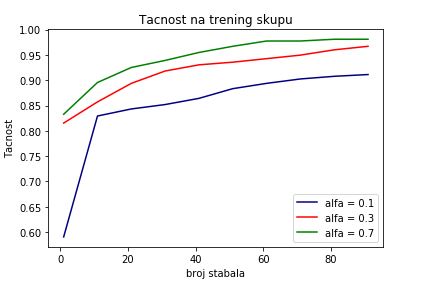


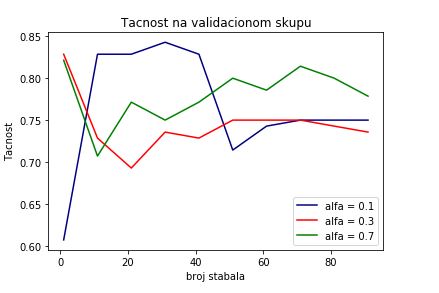


Dubina 5 ostvaruje najbolje rezultate na trening skupu, ali se vidi efekat preobučavanja. Ograničavanjem dubine stable na 4 eliminišemo preobučavanje I dobijamo bolju generalizaciju, dok smanjenjem dubine na dva nemamo sjajne rezultate ni na trening skupu.

1. Gradient Boosting – ispitivanje stope učenja u zavisnosti od veličine ansambla.

Za stopu učenja su isprobane vrednosti 0.1, 0.2, 0.7 i dobijeni su sledeći rezultati (kod u fajlu GB\_alfa.py):

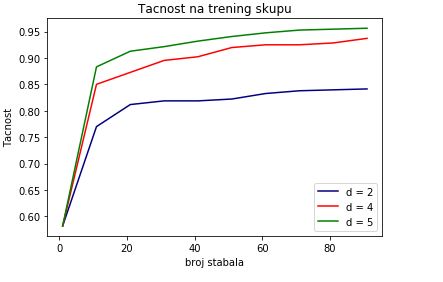


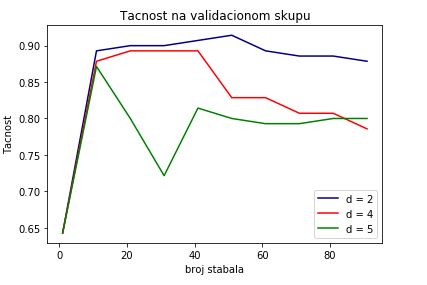


Sa grafika je jasno da što je stopa učenja manja, model sporije uči ali bolje generalizuje. Velike stope obučavanje dovode do preobučavanja na trening skupu.

1. Gradient Boosting – ispitivanje maksimalne dubine stable u zavisnosti od veličine ansambla.

Ponovo su ispitivane dubine 2, 4 i 5 I dobijeni su sledeći rezultati (kod u fajlu GB\_depth.py):





Za dubinu 5 je očigledno preobučavanje modela. Dubina 4 daje zadovoljavajuće rezultate do veličine ansambla oko 40 stabala, posle čega se preobučava, dok ograničavanjem dubine stabla na 2 dobijamo model koji lepo generalizuje.