

Analiza skupa podataka

“Classify gestures by reading muscle activity”

Seminarski rad u okviru kursa
Istraživanje podataka
Matematički fakultet

Aleksandra Stojanović
48/2016

18. maj 2019.

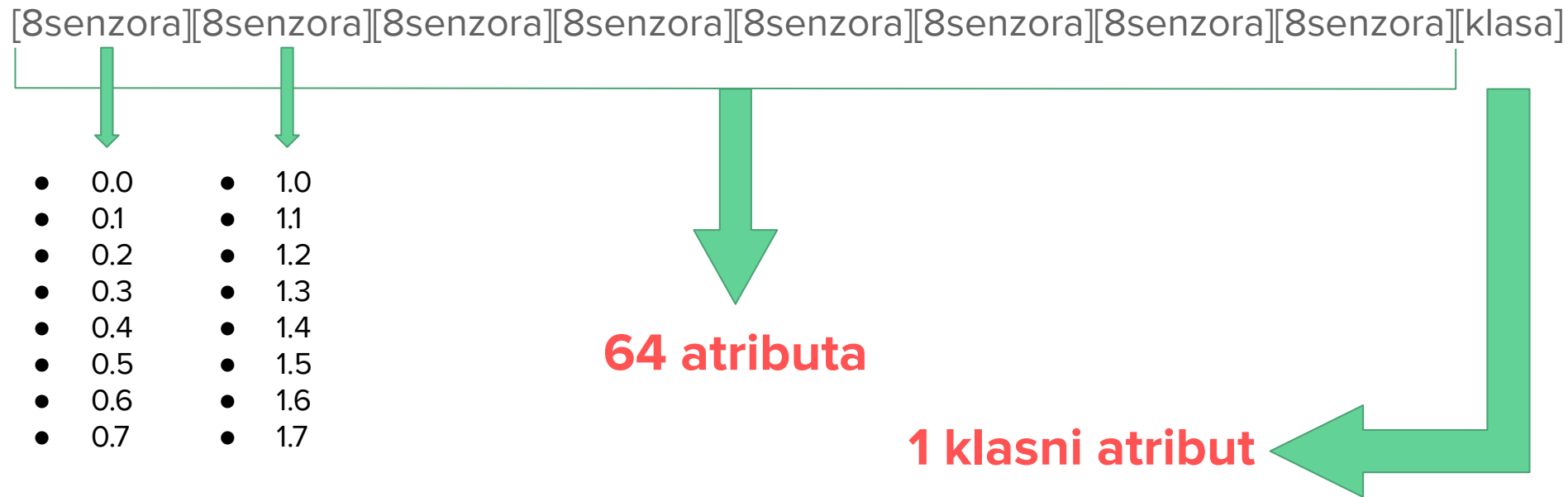
Opis skupa podataka

4 datoteke:

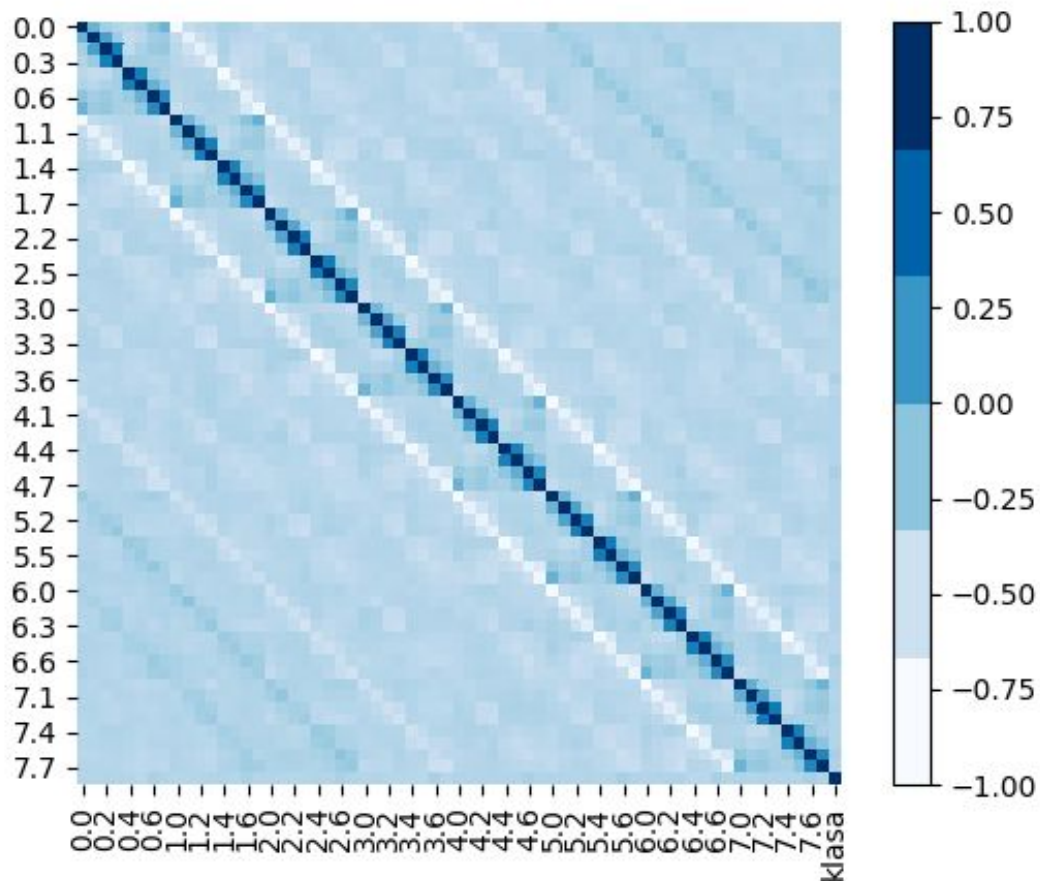
● 0.csv - kamen	2909
● 1.csv - makaze	2902
● 2.csv - papir	2942
● 3.csv - ostalo	2921
	<hr/>
	11674



Opis skupa podataka

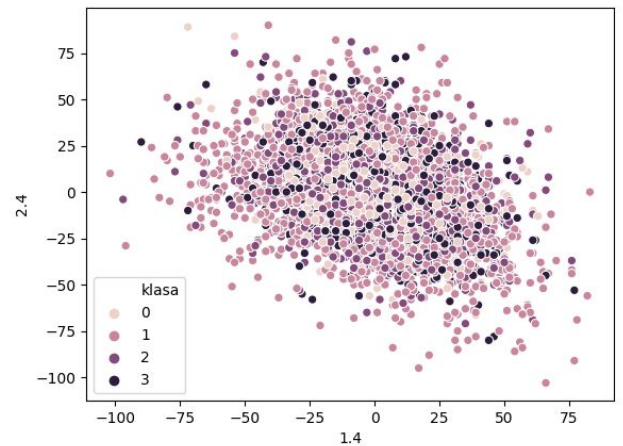
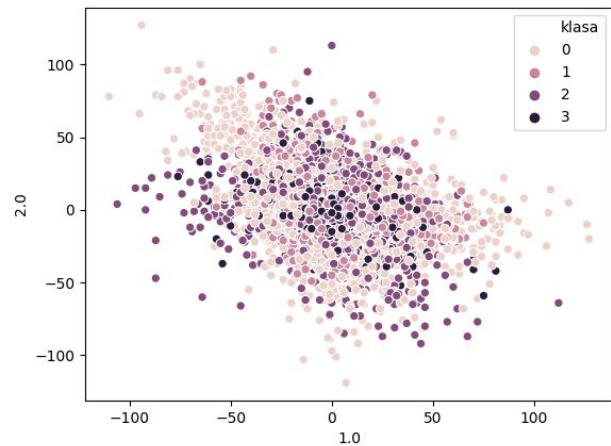
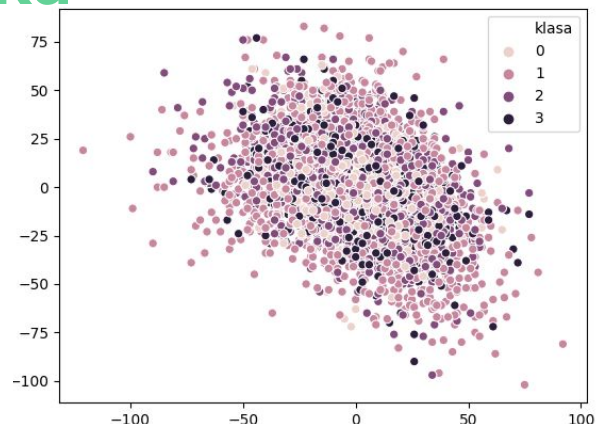
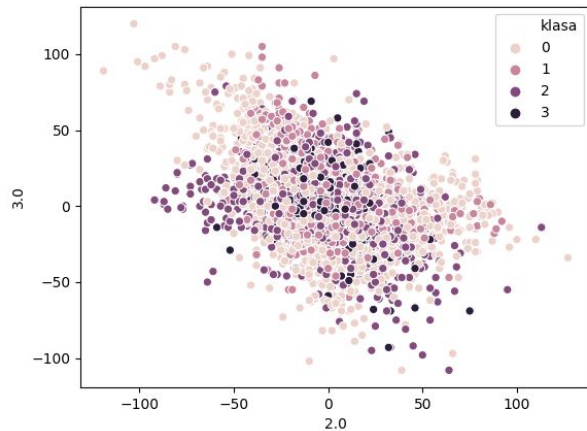
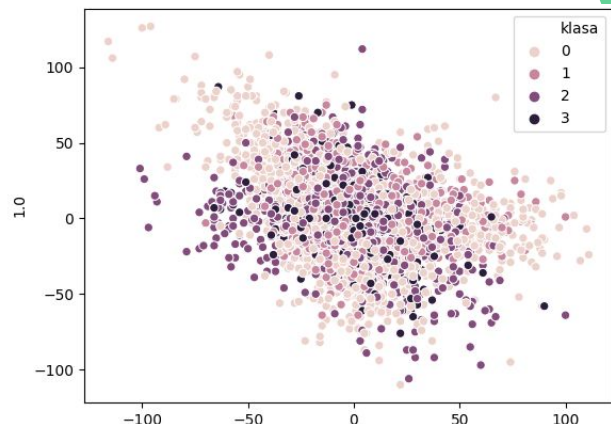


Vizualizacija podataka



- niska korelacija podataka
- intervali $[-0.6, 0.4]$
- nedovoljno pozitivno/negativno korelisani za smanjenje broja atributa

Vizualizacija podataka



C5.0

1_Training	Kamen	Makaze	Ostalo	Papir
Kamen	1979	4	32	27
Makaze	1	1925	78	44
Ostalo	67	79	1810	49
Papir	41	36	61	1915

2_Testing	Kamen	Makaze	Ostalo	Papir
Kamen	774	3	49	41
Makaze	2	722	83	47
Ostalo	56	93	693	74
Papir	49	71	83	686

Partition	1_Training	2_Training
Correct	93.63%	81.54%
Wrong	6.37%	18.46%

Partition	1_Training	2_Training
Correct	87.1%	81.85%
Wrong	12.9%	18.15%



povećanje strogosti pri
potkresivanju na 100

Partition	1_Training	2_Training
Correct	97.16%	90.64%
Wrong	2.84%	9.36%



uključivanje boosting opcije

C&Rt

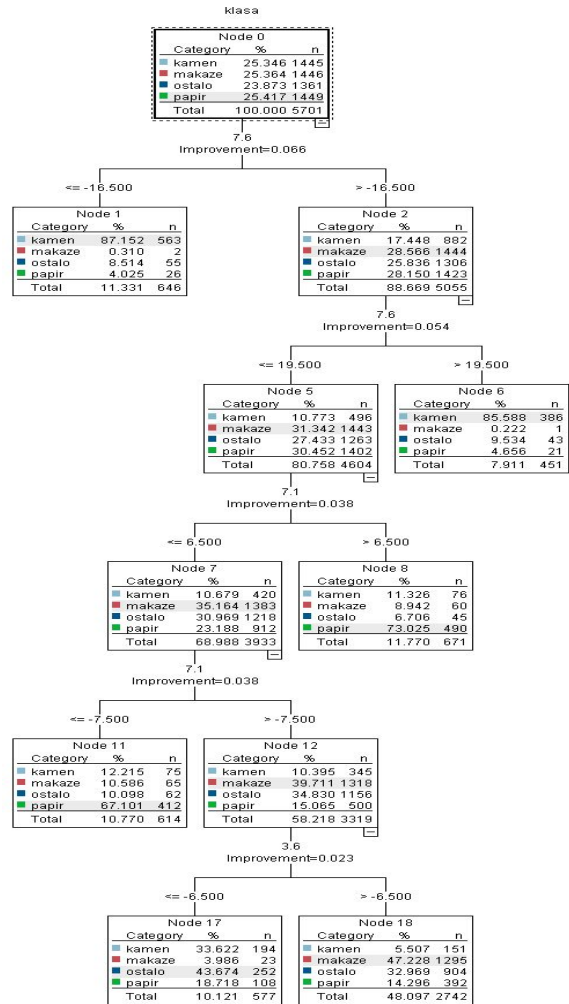
Partition5	1_Training	2_Training	Partition10	1_Training	2_Training
Correct	59.54%	58.88%	Correct	72.16%	71.38%
Wrong	40.46%	41.12%	Wrong	27.84%	28.62%

bez boosting opcije

Partition5	1_Training	2_Training
Correct	89.6%	85.99%
Wrong	10.4%	14.01%

Partition10	1_Training	2_Training
Correct	89.29%	86.98%
Wrong	10.71%	13.02%

sa boosting opcijom



KNN

Partition	1_Training	2_Training
Correct	76.26%	68.07%
Wrong	23.74%	31.93%



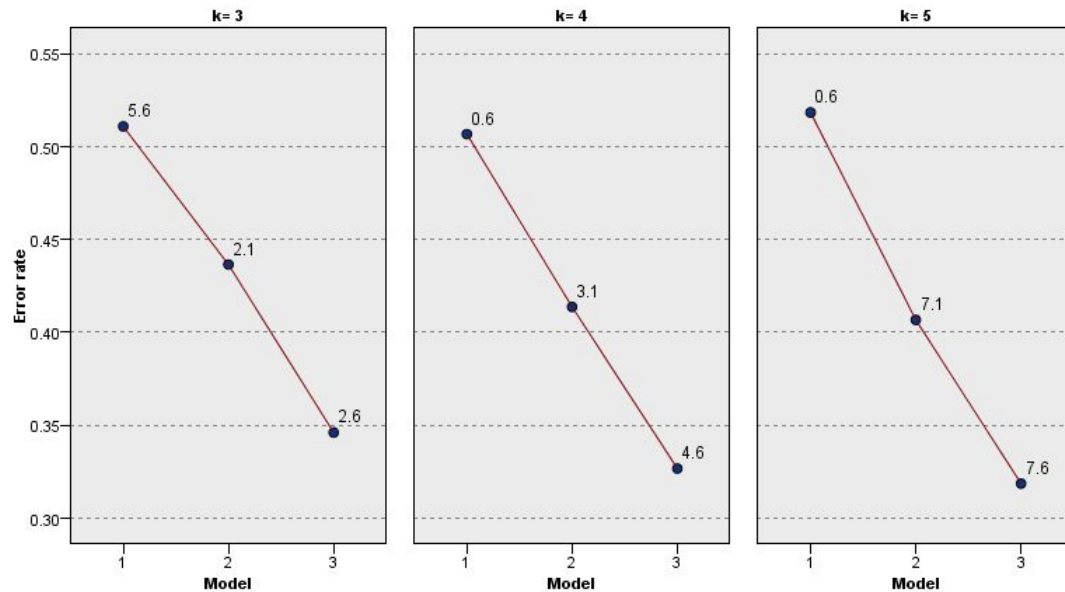
bez boosting opcije

Partition	1_Training	2_Training
Correct	74.83%	70.28%
Wrong	25.17%	29.72%



sa boosting opcijom

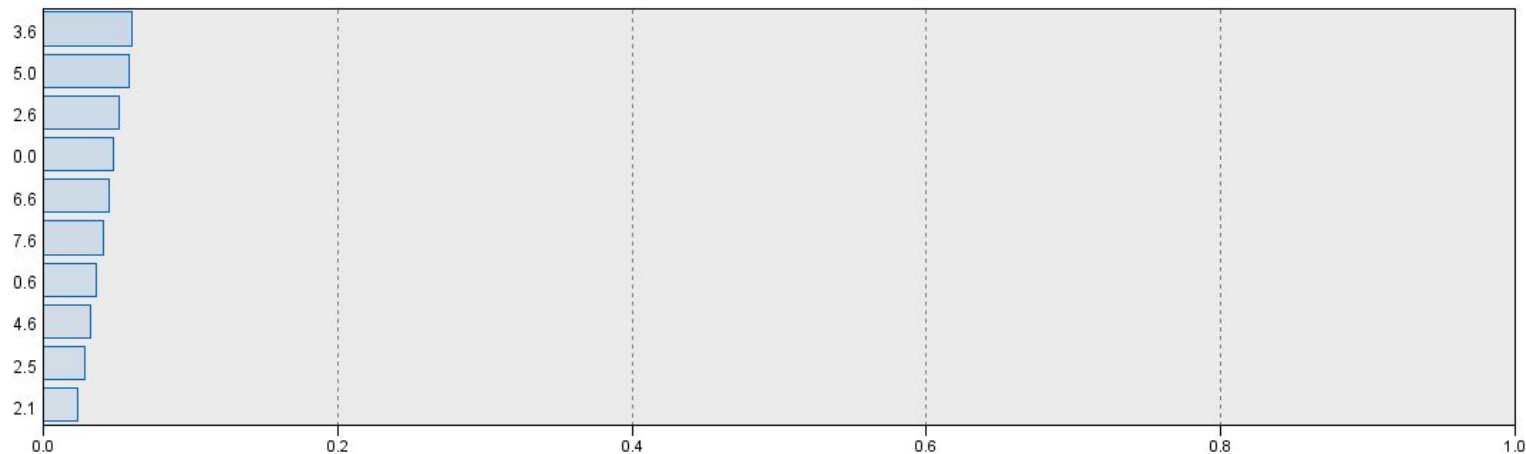
k and Predictor Selection



SVM

Predictor Importance

Target: klasa



Least Important

Partition	1_Training	2_Training
Correct	76.36%	74.31%
Wrong	23.64%	25.69%

Most Important

Partition	1_Training	2_Training
Correct	65.77%	68.03%
Wrong	34.23%	31.97%



PCA/Factor 50

Partition	1_Training	2_Training
Correct	68.26%	70.26%
Wrong	31.74%	29.74%



PCA/Factor 60

Hvala na pažnji!