mozgalo
gmolazo

Detecting Packed Executable Files

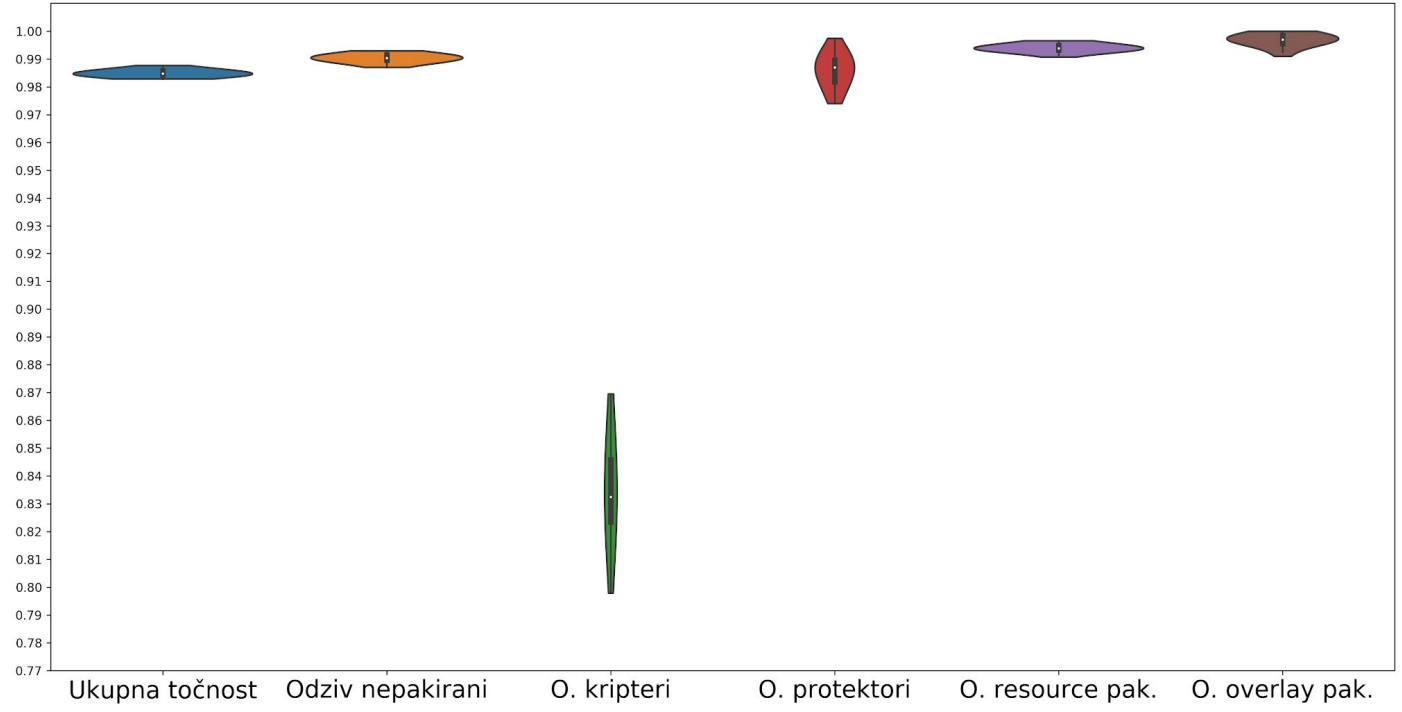
Metodologija

- odabir značajki
- odabir modela

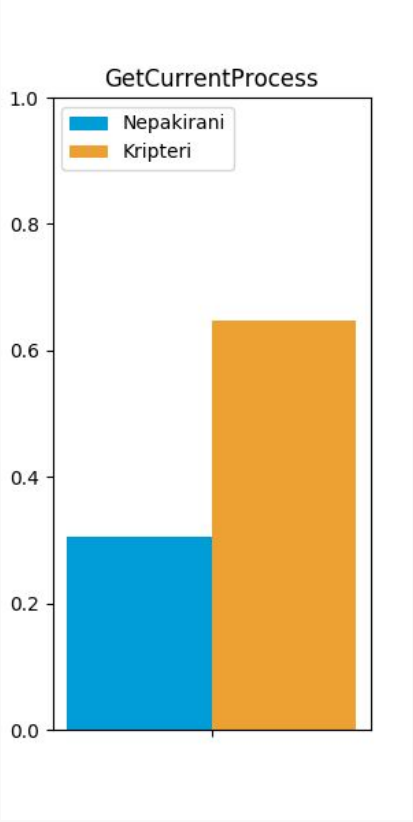
Traženje značajki

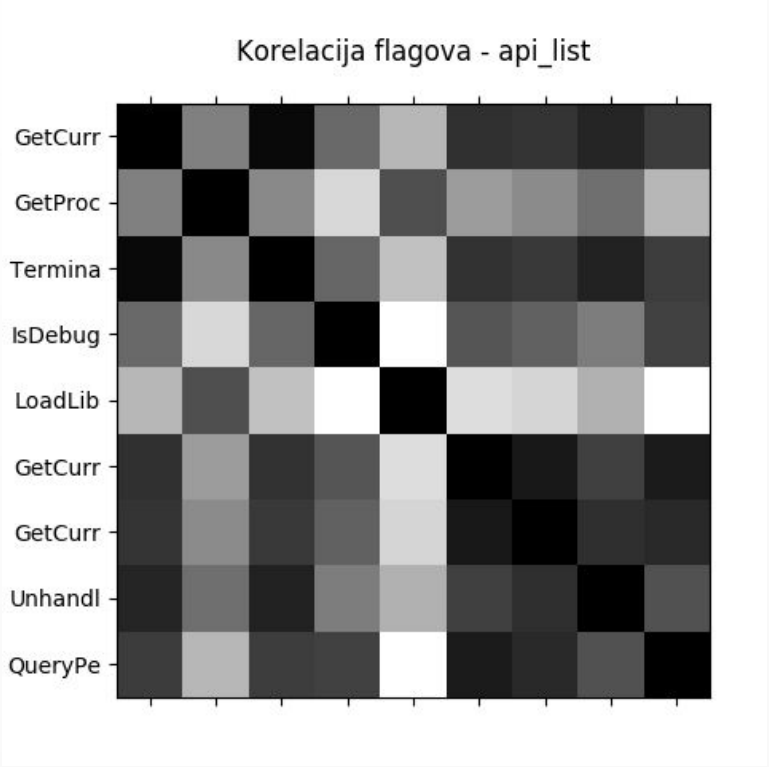
- radovi i radionice
- konzultacije
- vizualizacija

Značajke iz izvještaja

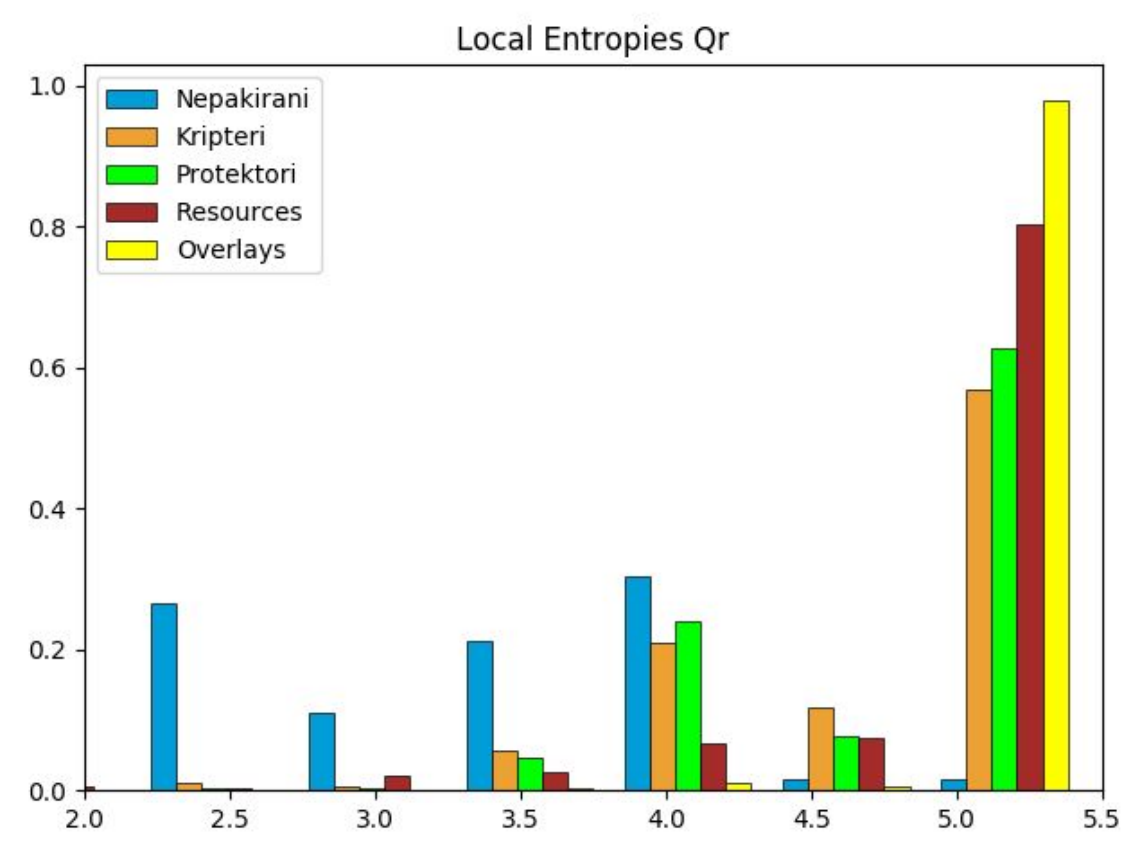


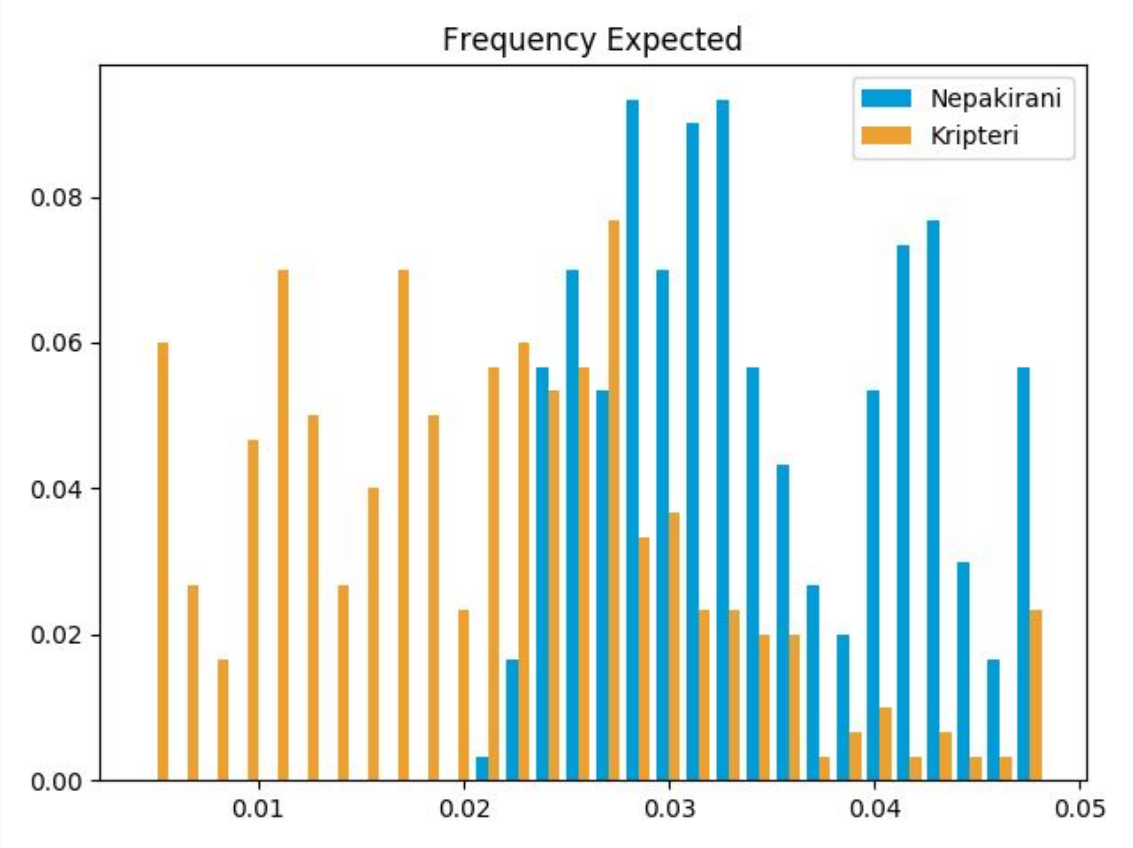
Kopanje po izvještajima





Značajke iz PE datoteka

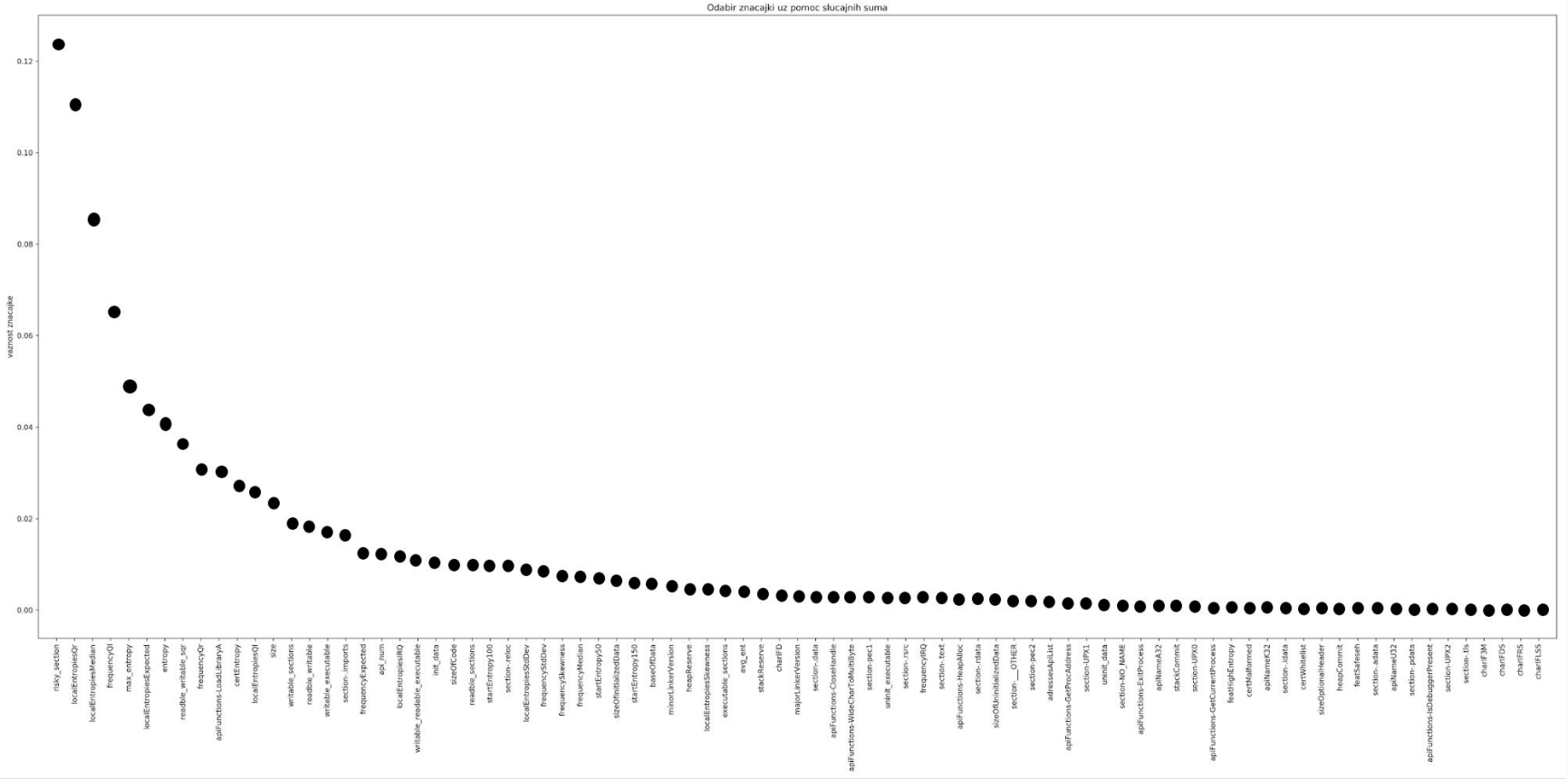



Frequency Expected
Nepakirani
Kripteri
0.08
0.06
0.04
0.02
0.00
0.01
0.02
0.03
0.04
0.05

Odabir modela

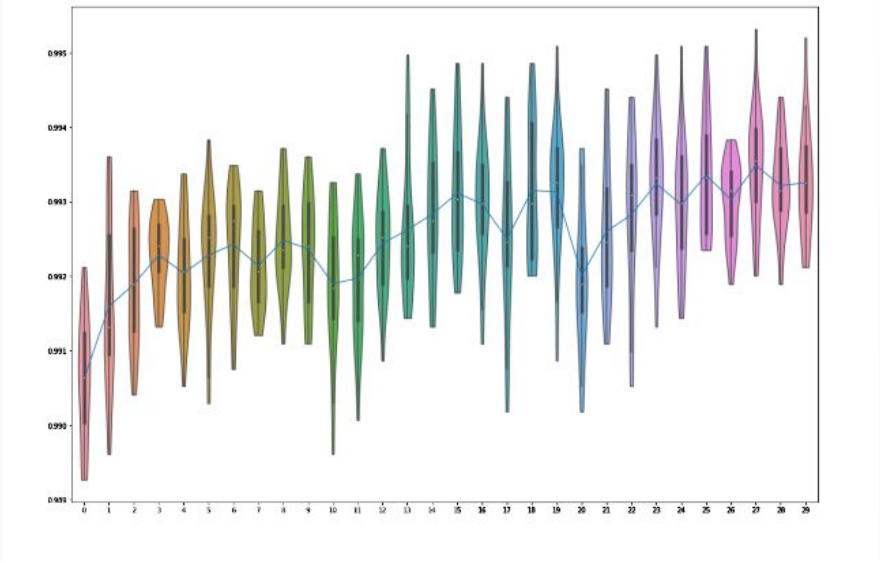
- Random Forest Classifier
  - brzo treniranje i evaluacija
  - kategoričke i numeričke varijable

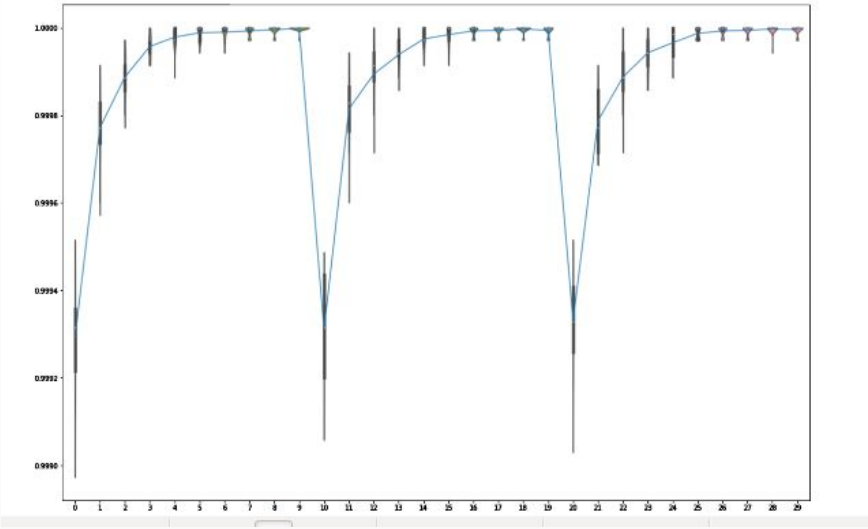
Važnost značajki po RFC-u



Odabir hiperparametara

- Distribucija točnosti ;unakrsna validacija
- Hiperparametri: broj stabala; max značajki po stablu; max dubina stabla

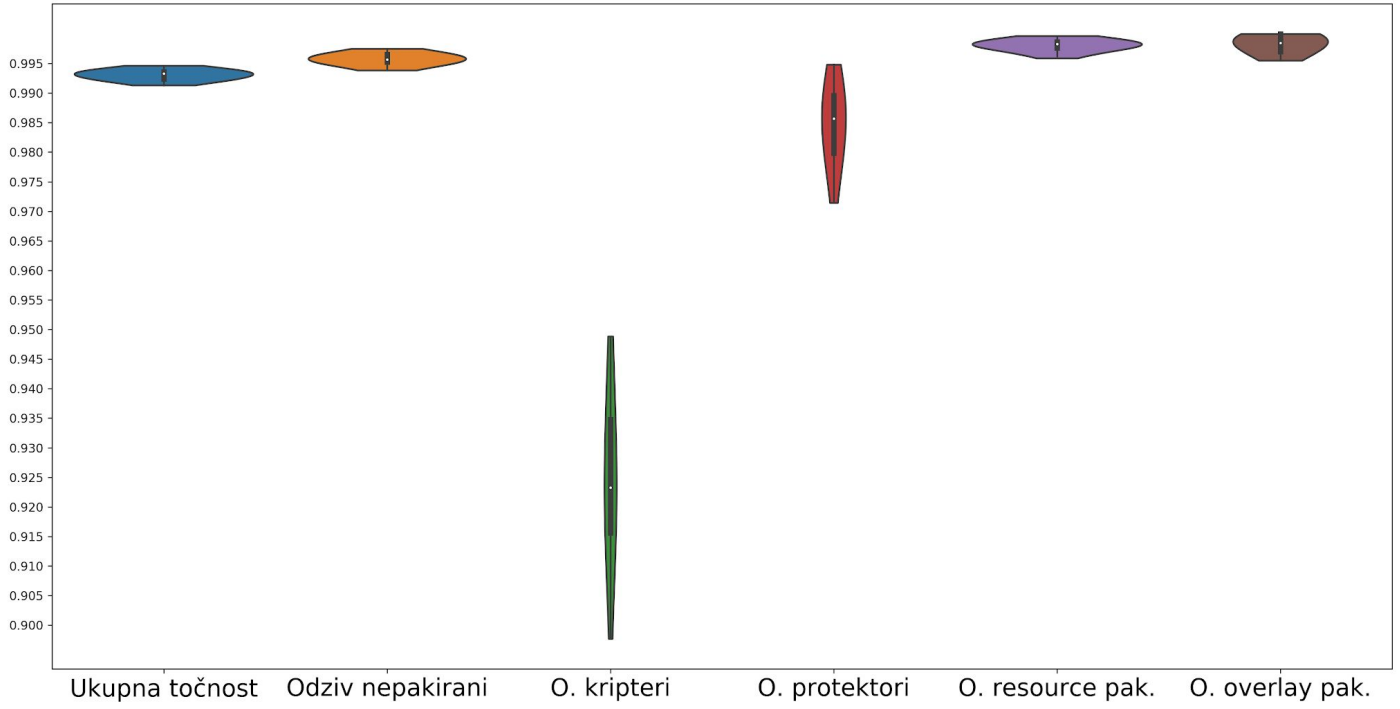




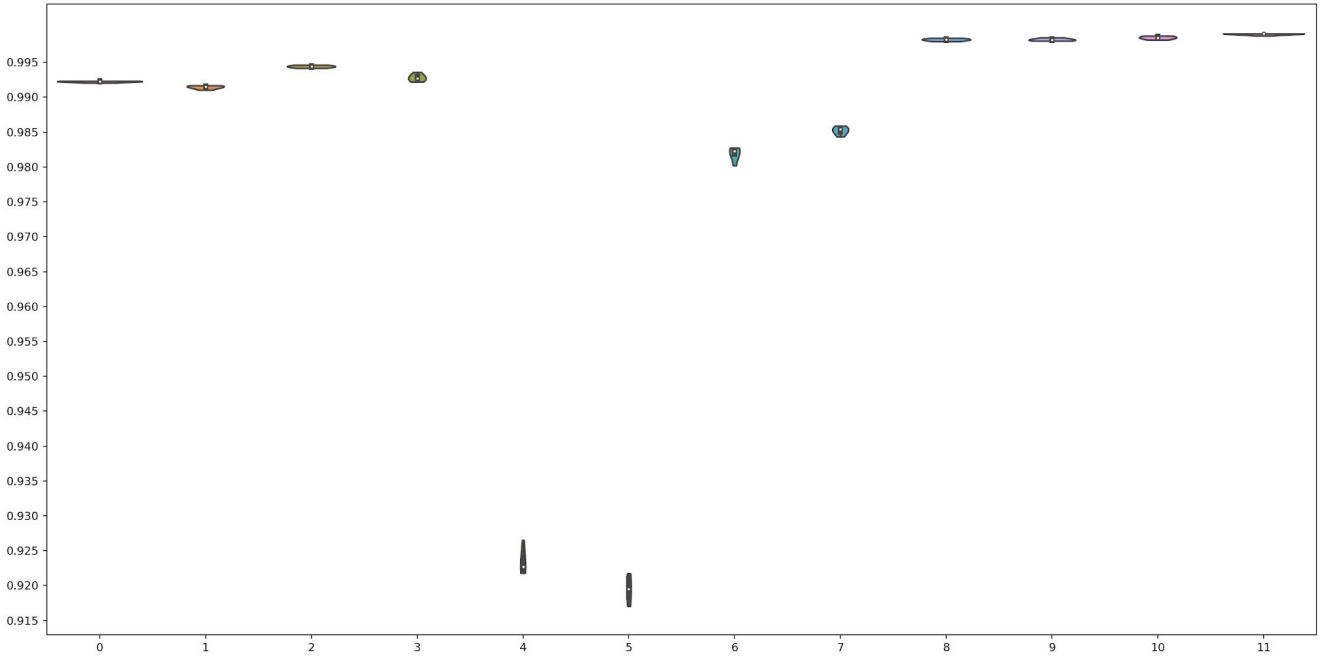
Rezultati

- 60 značajki
- Procjena točnosti: 99.2%

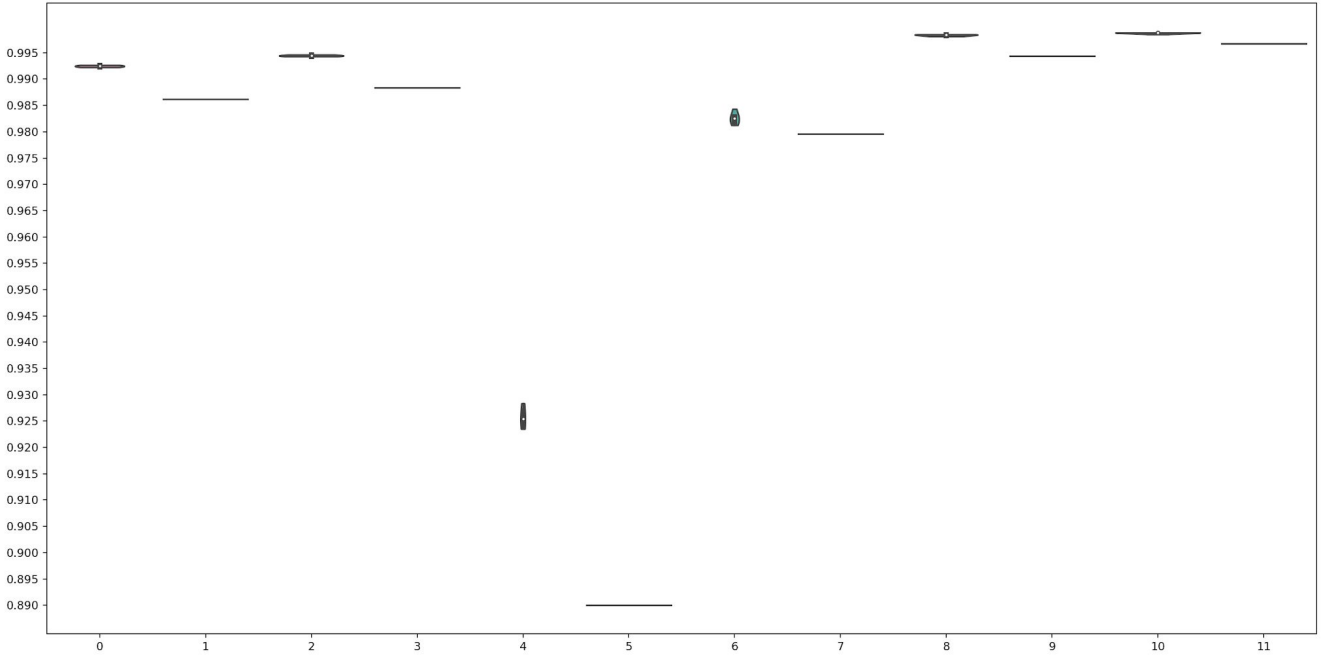
Rezultati


0.995
0.990
0.985
0.980
0.975
0.970
0.965
0.960
0.955
0.950
0.945
0.940
0.935
0.930
0.925
0.920
0.915
0.910
0.905
0.900
Ukupna točnost
Odziv nepakirani
O. kripteri
O. protektori
O. resource pak.
O. overlay pak.

Rezultati - sa vlastitim značajkama



Rezultati-RF vs AdaBoost



Zaključak

- izvještaji daju većinu podataka potrebnih za detekciju pakiranih datoteka
- naše značajke su korisne