

# Mašinsko učenje - predstavljanje sadržaja

Tatjana Jakšić Krüger

[tatjana@turing.mi.sanu.ac.rs](mailto:tatjana@turing.mi.sanu.ac.rs)

# Ciljevi kursa

- 1 Osposobljavanje studenta za razumevanje osnovnih metoda mašinskog učenja.
- 2 Upoznavanje sa osnovnom terminologijom zastupljenom u industrijskoj i akademskoj zajednici.
- 3 Priprema za metode mašinskog učenja koje se koriste za rešavanje prostornih problema (*eng.* spatial problems).

# Sadržaj

- Definicija, osnovni pojmovi i podele.
- Nadgledano i nenadgledano učenje, .
- Generalizacija i prilagođavanje, Okamova oštrica.
- Probabilistički modeli (linearna i logistička regresija, naivni Bajes).
- Metoda potpornih vektora.
- Ansambli.
- Evaluacija i izbor modela.
- Regularizacija.
- Optimizacija.
- Neuronske mreže i duboko učenje.

## Literatura i online predavanja



- A. Geron: Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd edition, O'Reilly, 2019.
- Stanford CS229 kurs: <http://cs229.stanford.edu/>
- MATF Uni Beograd: <http://ml.matf.bg.ac.rs/>

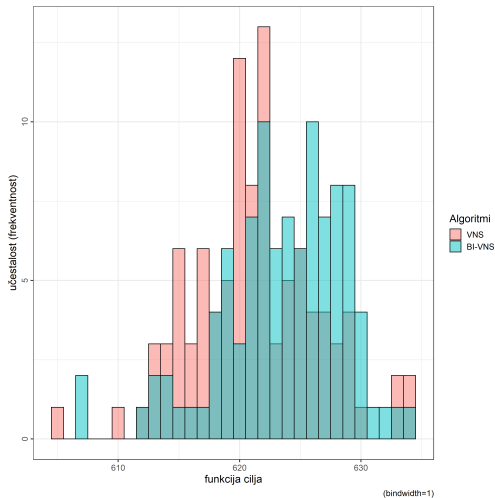
# Predznanje

- Linearna algebra:
  - množenje matrica,
  - operacije nad matricama (transpozicija, inverz, trag),
  - sopstvene vrednosti i sopstveni vektori.
- Teorija verovatnoće:
  - klasična i
  - bajesova verovatnoća.
- Statistika:
  - deskriptivna statistika,
  - raspodela podataka i njihova vizualizacija (funkcija gustine verovatnoće),
  - multivarijaciona statistička analiza (regresione metode).

# Vizualizacija podataka

□ Izdvajamo:

- Histogram
- Box-plot dijagram.
- Kumulativni dijagram.



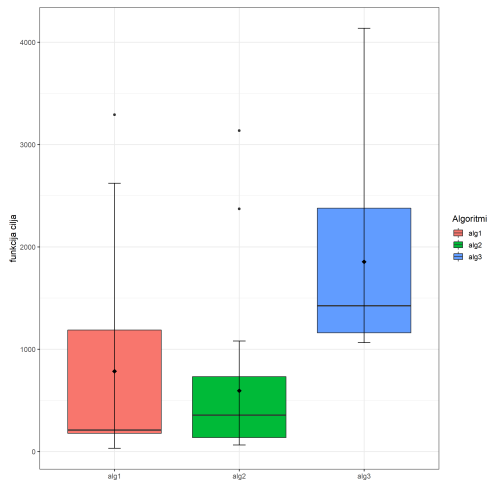
# Vizualizacija podataka

## □ Izdvajamo:

□ Histogram

□ **Box-plot dijagram.**

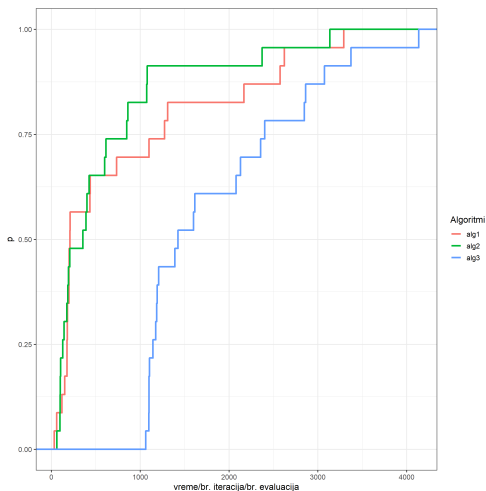
□ Kumulativni  
dijagram.



# Vizualizacija podataka

□ Izdvajamo:

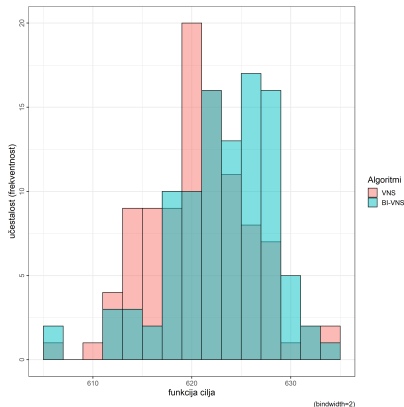
- Histogram
- Box-plot dijagram.
- Kumulativni dijagram.





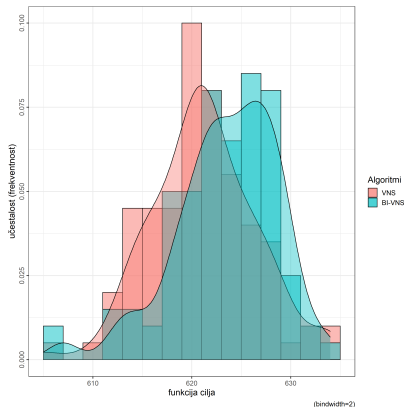
# Histogram - podsetnik

- Prednost:
  - Detekcija vrednosti oko koje se grupiše najviše podataka.
  - Jasan prikaz raspona vrednosti.
- Nedostaci:
  - Neophodno je namestiti veličinu binova.



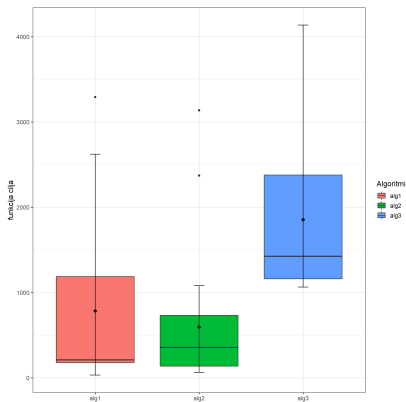
# Histogram - podsetnik

- Prednost:
  - Detekcija vrednosti oko koje se grupiše najviše podataka.
  - Jasan prikaz raspona vrednosti.
- Nedostaci:
  - Neophodno je namestiti veličinu binova.
- Empirijska kriva gustine.
  - Bolja od histograma da se utvrdi oblik krive raspodele.



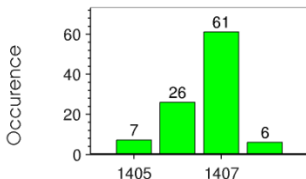
# Box-plot dijagram

- Pogodan da se prikaže:
  - raspon vrednosti;
  - odnos između srednje i medijana vrednosti;
  - uporedi više od dva uzoraka (algoritama).



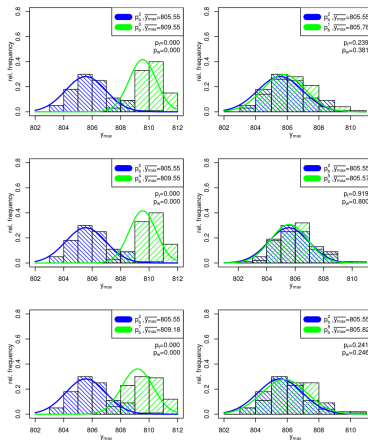
# Histogram

- Najbolji prvi korak kod vizualizacije uzorka.
- Pogodan za:
  - prebrojavanje;
  - pomaže opisu obliku krive raspodele uzorka;
  - poređenje dva uzorka (algoritma).



# Histogram

- Najbolji prvi korak kod vizualizacije uzorka.
- Pogodan za:
  - prebrojavanje;
  - pomaže opisu obliku krive raspodele uzorka;
  - poređenje dva uzorka (algoritma).



- Statistiku primenjujemo:
  - kada je broj grafika isuviše veliki da bismo izvukli smislene rezultate.
  - da bismo izbegli subjektivne zaključke putem grafika.
  - da bismo ispitali povezanost parametara (features).

# Naredna predavanja

- Definicija, osnovni pojmovi i podele.
- Nadgledano i nenadgledano učeje.



Hvala na pažnji.  
Da li imate pitanja?



# Ribbon

Convergence graphic for 100 different seeds. vns\_swapFI-LS, ogra50\_10

