



# Mašinsko učenje 2018

Zadatak 1



# Sadržaj

- Podsetnik
- Zadatak 1
- SciPy Stack
- Uputstvo



# Podsetnik



# Podsetnik

- Maksimalno 60 bodova.
- 6 domaćih zadataka i projekat.
- Timski rad - svaki član tima mora dati svoj doprinos, a bodovi dodeljeni članovima istog tima mogu da se razlikuju.
- Opcije:
  - Projekat = max 25 bodova
  - 3 zadatka (max 25) + projekat (max 25) = max 50 bodova
  - 4 i više zadataka (max 35) + projekat (max 25) = max 60 bodova
  - Nagrada za najuspešnije = max 60 bodova



# Podsetnik

- Zadaci:
  - 19.03. - 25.03. Jednostruka linearna regresija
  - 26.03. - 11.04. Višestruka regresija
  - 16.04. - 22.04. SVM
  - 23.04. - 04.05. Ansambl klasifikatora
  - 07.05. - 13.05. Klasterovanje
  - 14.05. - 20.05. PCA.



# Zadatak 1

- Jednostruka linearna regresija:
  - Prediktovati masu mozga (g) u zavisnosti od zapremine glave (cm<sup>3</sup>).
  - Zadatak je uspešno urađen ukoliko se na testnom skupu podataka dobije RMSE (Root Mean Square Error) manji od 87.
  - Rok: **25.03.2018. u 21:59h.**
  - Dostupne biblioteke: **NumPy i Pandas.**



# Zadatak 1

- Jednostruka linearna regresija (predavanje 2a):
  - Gradient Descent (Batch vs Stochastic)
  - Normal Equation
  - High-leverage points
  - Normalizacija:
    - Min-max
    - Z-score normalization
  - RMSE:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Predicted_i - Actual_i)^2}{N}}$$



# Scipy Stack





# Scipy Stack

- Za izradu zadataka koristiti **Python 3.5.x**.
- SciPy Stack:
  - NumPy
  - SciPy
  - Matplotlib
  - Ipython
  - Jupyter
  - Pandas
  - Sympy
  - Nose.



# Scipy Stack

- Za prvi zadatak, na platformi su instalirane biblioteke:
  - NumPy:
    - [Docs](#)
    - [Stanford Tutorial](#)
  - Pandas:
    - [Docs](#)
    - [Tutorials](#)
    - [10 Minutes to pandas](#)
- Za potrebe vizualizacije podataka i pisanje propratnog izveštaja od pomoći može biti biblioteka Matplotlib:
  - [Docs](#)
  - [Tutorial](#)



# Uputstvo



# Uputstvo

- Preporuka za implementaciju:
  - Metoda **fit(x, y)** za “fitovanje” trening podataka
  - Metoda **predict(x)** za predikciju vrednosti testnog skupa.
  - Metoda **calculate\_rmse(y\_true, y\_predict)** za računanje RMSE.
- Uputstvo za submit na platformi i propratni izveštaj:
  - [Enastava/MU-SIIT-18/Files/Vežbe/Uputstvo.pdf](#)