



**NAPOVEDOVANJE IZIDOV TEKEM  
BADMINTONA NA OSNOVI  
ANALIZE OMREŽIJ**

**Matjaž Hribernik**

**Mentor:izr. prof. dr. Marko Bajec**

# UVOD

- Vse večje zbirke podatkov dostopne na spletu
- Napovedovanje rezultatov
- Analiza omrežij



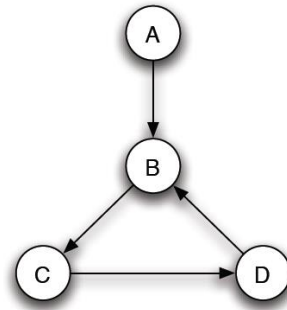
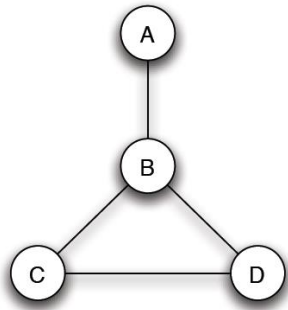
# BADMINTON

- Množica podatkov
  - A in B turnirji od sezone 2010/2011 do 2013/2014
- Način tekmovanja
  - Modificiran KO sistem
  - Glavni del in tekme za vsako mesto
- Pridobivanje podatkov
  - [www.tournamentsoftware.com](http://www.tournamentsoftware.com)



# OMREŽJA (1/2)

- Skupek vozlišč in povezav
- Usmerjena in neusmerjena omrežja



# OMREŽJA (2/2)

- Mere središčnosti
  - Mere središčnosti glede na dostopnost omrežja
  - Mere središčnosti glede na vmesnost omrežja
  - Mere središčnosti glede na stopnjo omrežja
    - Stopnja vozlišča
    - PageRank algoritem



# PODATKOVNO RUDARJENJE

- Odkrivanje vzorcev v velikih podatkovnih množicah s pomočjo:
  - Umetne inteligence
  - Strojnega učenja
  - Statistike



# ANALIZA TURNIRJA

- Pregled delovanja mer središčnosti na posameznem turnirju
  - Stopnja vozlišča
  - PageRank
  - Obteženi PageRank



# NAPOVED DOSEŽENEGA MESTA

- Napoved posameznih atributov
  - Najbolje napove atribut Število turnirjev
- Število turnirjev, InDeg, OutDeg in Uteženi PageRank

Metoda	MAE	RAE
Random Forest	7.5160	0.2163
Regression Tree	8.3673	0.2408
SVM Regression	31.5757	0.9087
kNN	6.3888	0.1839
Linear Regression	19.5630	0.5630
Mean	34.9311	1.0053





# NAPOVED TOČK

- Razlika med učinkovitostjo navadnega PageRank algoritma in obteženega PageRank algoritma
- Atributi Število turnirjev, InDeg, OutDeg in PageRank

Metoda	MAE	RAE
Random Forest	58.1575	0.1830
Regression Tree	77.1902	0.2430
SVM Regression	269.2776	0.8475
kNN	44.5108	0.1401
Linear Regression	23.0892	0.0727
Mean	319.6028	1.0059

- Atributi Število turnirjev, InDeg, OutDeg in obteženi PageRank

Metoda	MAE	RAE
Random Forest	54.9678	0.1730
Regression Tree	74.1638	0.2334
SVM Regression	268.8273	0.8461
kNN	40.4351	0.1273
Linear Regression	19.7204	0.0621
Mean	319.6028	1.0059



# NAPOVED POVPREČNEGA ŠTEVILA TOČK IGRALCA NA POSAMEZEN TURNIR

- Absolutne vrednosti posameznih atributov zamenjajo povprečne vrednosti glede na posamezen turnir

Metoda	MAE	RAE
Random Forest	6.2764	0.6931
Regression Tree	7.8542	0.8673
SVM Regression	7.3034	0.8065
kNN	6.3047	0.6962
Linear Regression	5.4203	0.5985
Mean	9.1072	1.0057



# PRIMERJAVA REZULTATOV IN ZAKLJUČEK

## ○ Analiza rezultatov

- Velikost podatkovne množice
- Način tekmovanja
- Obteženi PageRank algoritem

## ○ Možne izboljšave

- Večja in bolje porazdeljena podatkovna množica
- Časovna komponenta



○ **HVALA ZA POZORNOST**

