Fakulteta za Računalništvo in Informatiko

Večna pot 113, Ljubljana

1. Seminarska naloga pri predmetu UMETNA INTELIGENCA

Prof. dr. Igor Kononenko  
as. dr. Peter Vračar

Avtorja:

Simon Korošec, 63160171

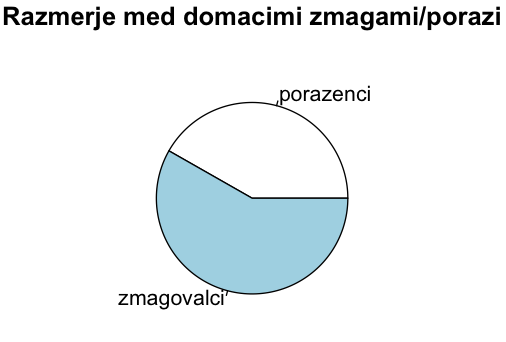
Gašper Smerkolj, 63160285

Priprava podatkov

Iz vseh podatkov sva najprej določila, če je domača ekipa zmagala tako da sva odštela od atributa HPTS atribut APTS. S tem sva lahko določila točnost najinega modela za zmage domače ekipe. Dodala sva še atribute kot so povprečje zadnjih 5 tekem, povprečje uspešnih metov za 3 v zadnjih 5 tekmah, ter povprečje skokov v obrambi za zadnjih 5 tekem. Te atribute sva naknadno dodala po ocenjevanju atributov. Potem sva odstranila atribute SEASON, DATE, HPTS, APTS, HOME, AWAY, saj bi naju pri klasifikaciji ter regresiji ovirali s tem da bi model ucil na nepotrebnih podatkih.

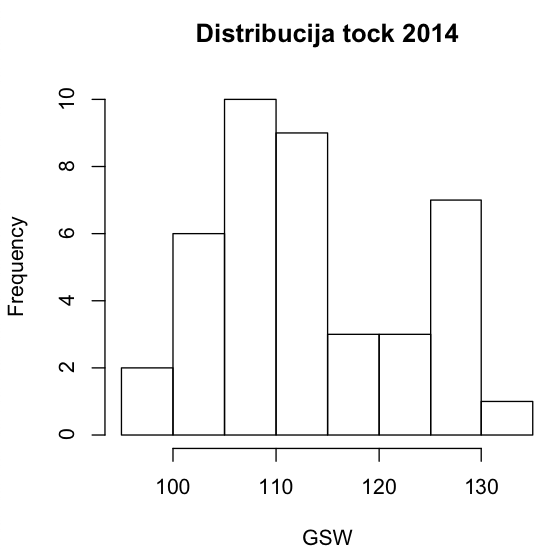
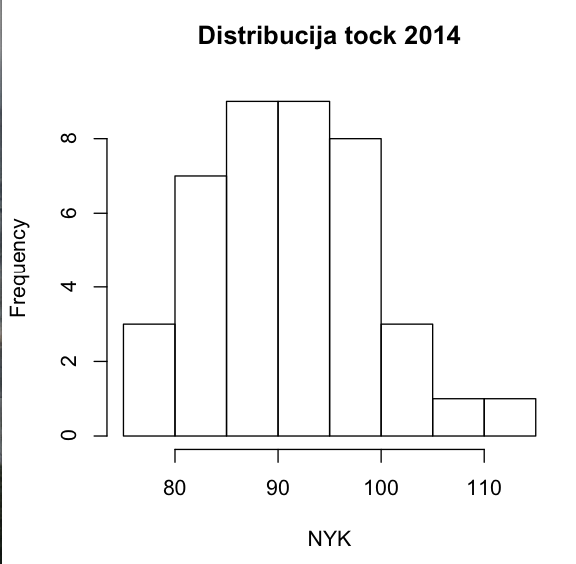
Vizualizacija podatkov

Prvi graf nam prikazuje zmage domače ekipe, ter poraze. Kot vidimo je v skupnih treh skupinah več ekip zmagalo doma kot pa zgubilo.

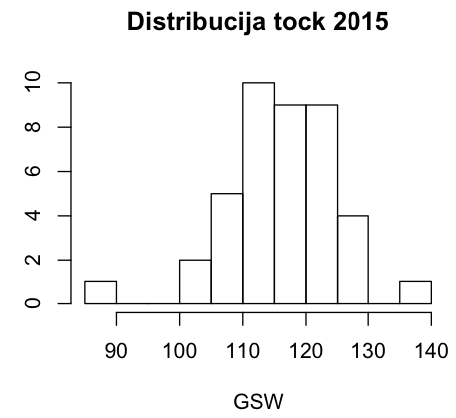
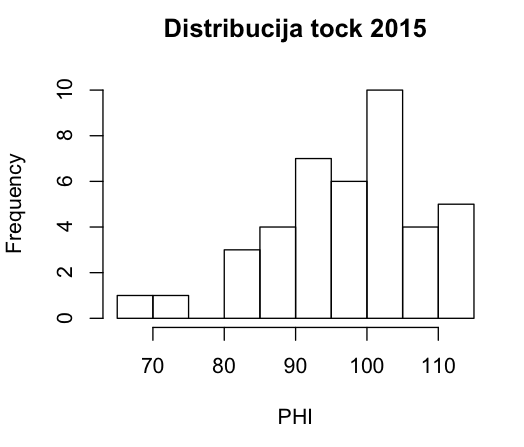


Naslednji grafi nam prikazujejo distrubucijo točk med ekipami ki so dosegle največ zmag, ter na desni ekipe z najmanj zmagami, za vsako sezono, ki so nam bile podane.

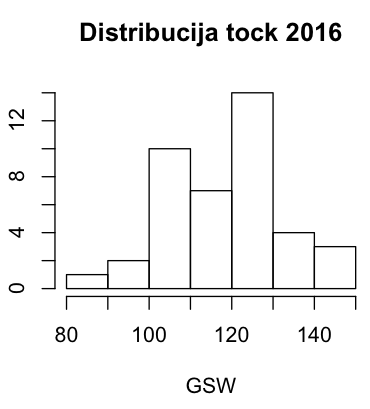
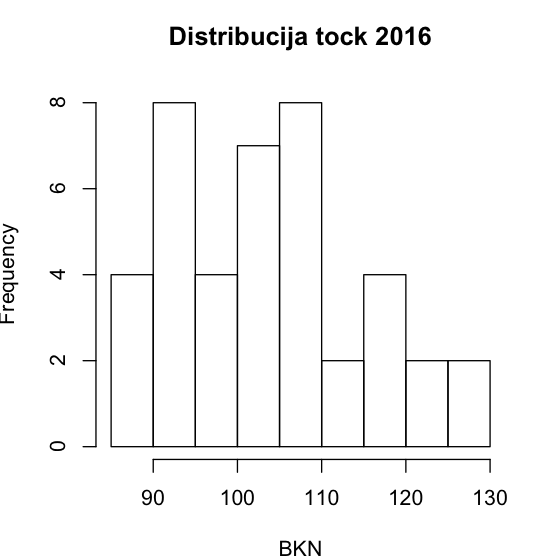
* Sezona 2014/2015

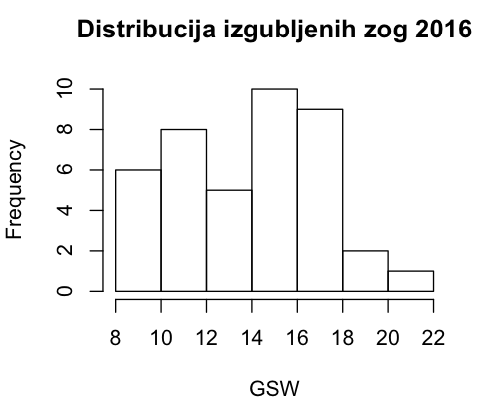
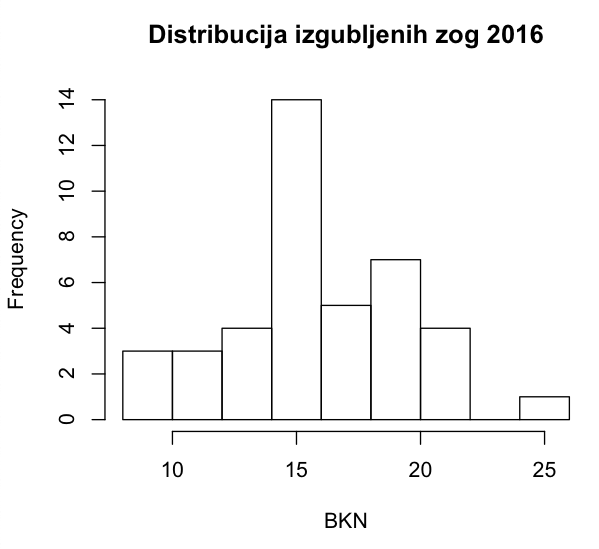
* Sezona 2015/16

* Sezona 2016/17

Naslednja grafa prikazujeta distribucijo izgubljenih žog med levo ekipo z največ zmagami, ter desno ekipo z najmanj zmagami, v sezoni 2016.

Regresija

Za regresijsko točnost sva izdelala več regresijskih formul

**1.formula**: Vse atribute, v že pripravljenih podatkih

**2.formula**:5 najboljših atributov, ki sva jih dobila z regresijskim ocenjevanjem

**3.formula**:7 najboljših atributov, ter še povprecje petih tekem ter, povprečje 5 tekem trojk

**4.formula** :7 najbolših atributov.

Ciljni spremenljivka je bila točke domačih minus točke gostujočih

S temi pristopi sva se lotila ocenjevanja z različnimi modeli kot so: linearni model, regresijsko drevo, naključni gozdovi, metoda podpornih vektorjev, kNN ter nevronska mreža.

* Formula1("rezultat\_razlika ~ HDRB + ADRB + ATOV + HTOV + H3PM")

Evaluacija: Pri formuli 1 so se najbolj odrezale nevronske mreže

|  |  |
| --- | --- |
| Formula 1 | |
| Metoda | Točnost |
| Linearni model | MAE:5.167; RMAE:0.488; MSE:42.44; RMSE:0.243 |
| Regresijsko drevo | MAE:8.39; RMAE:0.79; MSE:117.96; RMSE:0.67 |
| Naključni gozdovi | MAE:6.14; RMAE:0.58; MSE:61.86; RMSE:0.35 |
| SVM | MAE:5.25; RMAE:0.49; MSE:44.28; RMSE:0.254 |
|
| KNN | RMAE:0.68; |
| Nevronske mreže | MAE:5.168; RMAE:0.483; MSE:42.43; RMSE:0.244 |

* Formula2("rezultat\_razlika ~.,")

Evaluacija: Pri formuli 1 so se najbolj odrezale nevronske mreže

|  |  |
| --- | --- |
| Formula 2 | |
| Metoda | Točnost |
| Linearni model | Vsi podatki so bili pod 1.693611e-13 |
| Regresijsko drevo | MAE:8.38; RMAE:0.792; MSE:112.48; RMSE:0.646 |
| Naključni gozdovi | MAE:1.96; RMAE:0.18; MSE:8.42; RMSE:0.048 |
| SVM | MAE:5.25; RMAE:0.49; MSE:44.28; RMSE:0.254 |
|
| KNN | RMAE:0.63; |
| Nevronske mreže | MAE:0.11; RMAE:0.03; MSE:0.01; RMSE:0.0001 |

* Formula3("rezultat\_razlika ~ HAST + HDRB + ADRB +AAST + H3PM + A3PM + H2PM + povprecje5tekm + povprecjeH3PM")

|  |  |
| --- | --- |
| Formula 3 | |
| Metoda | Točnost |
| Linearni model | MAE:6.15; RMAE:0.58; MSE:58.8; RMSE:0.33 |
| Regresijsko drevo | MAE:8.65; RMAE:0.81; MSE:119.70; RMSE:0.68 |
| Naključni gozdovi | MAE:6.52; RMAE:0.61; MSE:66.74; RMSE:0.38 |
| SVM | MAE: 6.37; RMAE: 0.60; MSE: 64.72; RMSE: 0.37 |
|
| KNN | RMAE:0.684; |
| Nevronske mreže | MAE: 10.58; RMAE: 174.01; MSE:0.99; RMSE:1 |

* Formula4(“rezultat\_razlika ~ HAST + HDRB + ADRB +AAST + H3PM + A3PM + H2PM")

|  |  |
| --- | --- |
| Formula 4 | |
| Metoda | Točnost |
| Linearni model | MAE: 6.17; RMAE:0.58; MSE: 59.11; RMSE:0.34 |
| Regresijsko drevo | MAE:8.38; RMAE:0.79; MSE:112.48; RMSE:0.64 |
| Naključni gozdovi | MAE:6.50; RMAE:0.614; MSE:66.66; RMSE:0.38 |
| SVM | MAE: 6.29; RMAE: 0.59; MSE: 63.9; RMSE: 0.36 |
|
| KNN | RMAE:0.689 |
| Nevronske mreže | MAE: 6.16; RMAE: 59.18; MSE:0.59; RMSE:0.34 |