

Predikcija pobjednika Eurosong-a

Kim Novak Mentor: Stefan Anđelić

Fakultet Tehničkih Nauka

Uvod

Izbor za pjesmu evrovizije je takmičenje u kom zemlje iz Evrope šalju svog predstavnika koji se takmiči sa predstavnicima ostalih zemalja.

Pobjednika bira publika SMS glasanjem.

Posljedica ovog načina glasanja je što stanovnici neke zemlje glasaju za predstavnika zemlje suseda ili bivše kolonije ili bivše državne zajednice.

Motivacija

Ispitati da li broj zemalja koje su susedi posmatranoj državi utiče na plasman te države.

U obzir će se uzeti i zemlje koje su određenu zemlju nagrađivale poenima u barem 50% posmatranih slučajeva.

Dataset

Posmatrani podaci mogu se preuzeti sa sledećeg link-a:

<https://www.kaggle.com/c/Eurovision2010/data?2010%20Data.csv>

Dataset sadrži podatke o državama, plasmanu, primljenim bodovima od svake države učesnice, regionu kom pripadaju, pesmi, izvođačima itd. za takmičenja koja su održana od 1998-2009. godine.

Model i metod

Metod kojim će se raditi predikcija je Linearna Regresija. Kao nezavisna promenljiva uzeće se broj prijateljskih zemalja dok će zavisna promenljiva biti rang tj. mesto koje je osvojila posmatrana država na takmičenju. Svaka država sadržaće podatak o broju prijateljskih država koje su učestvovala u istom polufinalu ili finalu kao i ona i koje mesto je zauzela te godine. Na osnovu toga odrediće se koeficijenti linearne regresije i utvrditi da li je plasman zaista zavisan od broja prijateljskih zemalja.

Rezultati, analiza i budući rad

Na osnovu 461 para testiranih (x,y) vrednosti polinom koji je dobijen fit-ovanjem vrednosti na polinom oblika $y=k*x+n$ ima vrednosti koeficijenata k i n -0.315 i 14.63, respektivno. Funkcija je opadajuća, odnosno za veće x sve je manje y, tj. što više komšija jedna država ima to bolji plasman zauzima. Međutim, kao što se može primetiti na slici ispod levo, ovaj faktor ne utiče dovoljno da bi država osvojila neko od prvih 9 mesta. Na slici ispod desno prikazana je predikcija vrednosti y za unetu vrednost x = 15.

Pošto se na osnovu rezultata može zaključiti da broj prijateljskih država jeste faktor koji utiče na plasman ali nedovoljno da bi država ostvarila neki visok plasman, uzeti u obzir druge faktore kao što je jezik na kom se izvodi pesma, da li je izvodi muškarac, žena ili bend i sl.

