

Predviđanje uspjeha bankarskog marketinga

Projektni prijedlog za kolegij Strojno učenje

Davorin Gradečak, Petra Jambriško, Jakov Krunić

21. travnja 2020.

Opis problema

Predviđanje
uspjeha
bankarskog
marketinga

Davorin Gradečak,
Petra Jambriško,
Jakov Krunić

Predviđanje uspjeha bankarskog marketinga na temelju popisnih podataka:

- ▶ *train* podaci - 41188 instanci/ispitanih korisnika
- ▶ *test* podaci - 4119 instanci/ispitanih korisnika

Svaki korisnik je opisan s 20 atributa i atributom *y* koji predviđamo (target).

Atribut *y* nam govori je li korisnik ugovorio oročeni depozit ili ne - *y* je kategorijska varijabla.

Opis i cilj problema

Analiza podataka

Metode, algoritmi i
ocjene njihove
uspješnosti

Očekivani rezultati
predloženog
projekta

Cilj projekta

Predviđanje
uspjeha
bankarskog
marketinga

Davorin Gradečak,
Petra Jambriško,
Jakov Krunić

Opis i cilj problema

Analiza podataka

Metode, algoritmi i
ocjene njihove
uspješnosti

Očekivani rezultati
predloženog
projekta

- ▶ odrediti utjecaj pojedinih atributa na y (target)
- ▶ napraviti što uspješniji klasifikator, tj. model koji što uspješnije predviđa hoće li korisnik ugovoriti oročeni depozit

Hipoteza: Moguće je primjenom adekvatnih algoritama strojnog učenja dobiti model koji uspješno predviđa odluku korisnika.

Analiza podataka

Predviđanje
uspjeha
bankarskog
marketinga

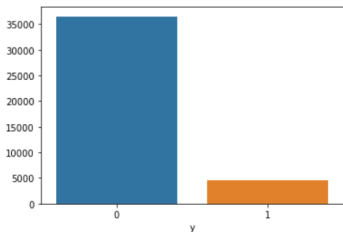
Davorin Gradečak,
Petra Jambriško,
Jakov Krunić

Opis i cilj problema

Analiza podataka

Metode, algoritmi i
ocjene njihove
uspješnosti

Očekivani rezultati
predloženog
projekta



U *train* podacima, od svih korisnika, njih 36548 nije ugovorilo oročeni depozit, a njih 4640 jest. Negativnih slučajeva je skoro 8 puta više od pozitivnih.

Analiza podataka

Predviđanje
uspjeha
bankarskog
marketinga

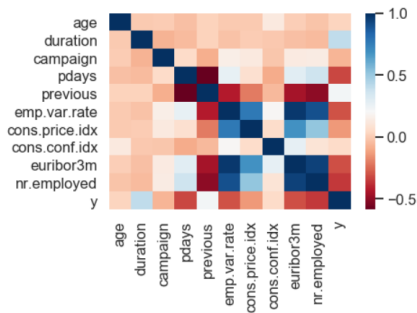
Davorin Gradečak,
Petra Jambriško,
Jakov Krunić

Opis i cilj problema

Analiza podataka

Metode, algoritmi i
ocjene njihove
uspješnosti

Očekivani rezultati
predloženog
projekta



Korelacijska
matrica atributa
pokazuje da je
target varijabla y
najviše korelirana s
trajanjem poziva.

Metode, algoritmi, tehnike i alati

Predviđanje
uspjeha
bankarskog
marketinga

Davorin Gradečak,
Petra Jambriško,
Jakov Krunić

Opis i cilj problema

Analiza podataka

Metode, algoritmi i
ocjene njihove
uspješnosti

Očekivani rezultati
predloženog
projekta

Algoritmi koje ćemo koristiti:

- ▶ logistička regresija
- ▶ naivni Bayes
- ▶ k -najbližih susjeda
- ▶ slučajne šume
- ▶ metoda potpunih vektora
- ▶ neuronske mreže

Metode za poboljšanja i uspoređivanja uspješnosti:

- ▶ preuređivanje podataka
- ▶ korištenje PCA (Analiza glavnih komponenata)
- ▶ način rješavanja problema nedostajućih vrijednosti

Ocjene uspješnosti metoda i algoritama

Predviđanje
uspjeha
bankarskog
marketinga

Davorin Gradečak,
Petra Jambriško,
Jakov Krunić

Opis i cilj problema

Analiza podataka

Metode, algoritmi i
ocjene njihove
uspješnosti

Očekivani rezultati
predloženog
projekta

- ▶ Evaluacija pomoću vrijednosti AUC (Area Under Curve), tj. površine ispod ROC (Receiver Operating Characteristic) krivulje
- ▶ Matrica konfuzije
- ▶ Vrijednost preciznosti P i F_1 , tj. mjera koja povezuje preciznost i osjetljivost

Očekivani rezultati predloženog projekta

Predviđanje
uspjeha
bankarskog
marketinga

Davorin Gradečak,
Petra Jambriško,
Jakov Krunić

Opis i cilj problema

Analiza podataka

Metode, algoritmi i
ocjene njihove
uspješnosti

Očekivani rezultati
predloženog
projekta

Kao konačni rezultat projekta očekujemo model koji što uspješnije rješava dani klasifikacijski problem.