

Sustav za preporuku

Ovaj sustav za preporuku koristi algoritam kolaborativnog filtriranja primijenjen na podacima o ocjenama piva prikupljenih sa stranice BeerAdvocate. Na temelju prikupljenih ocjena korisnika sa spomenute stranice, korisniku aplikacije nudimo da unese svoje ocjene te preporučujemo listu piva koja bi mu se mogla svidjeti.

Priprema podataka

Za dane ocjene, u matricu Y (broj piva \times broj korisnika) spremamo vrijednosti $y^{(i,j)}$ ($1 - 5$) koje predstavljaju ocjenu korisnika j za pivo i . Budući da nemamo ocjene svih korisnika za sva piva, potrebna nam je indikatorska matrica R , gdje je $R(i, j) = 1$ ako je korisnik j ocijenio pivo i , inače je $R(i, j) = 0$. Cilj kolaborativnog filtriranja je predvidjeti ocjene piva koja nisu ocijenjena (gdje je $R(i, j) = 0$) što će nam omogućiti preporuku piva koja imaju najvišu predviđenu ocjenu.

Također, potrebni su nam sljedeći vektori: vektor značajki $x^{(i)}$ za svako pivo i te vektor parametara $\theta^{(j)}$ za svakog korisnika j . Sve navedene vektore na početku inicijaliziramo na male proizvoljne vrijednosti.

Kolaborativno filtriranje

Kolaborativno filtriranje za preporuku piva uz dane parametre $x^{(1)}, \dots, x^{(n_b)}$ i $\theta^{(1)}, \dots, \theta^{(n_u)}$, predviđa ocjenu korisnika j za pivo i kao $y^{(i,j)} = (\theta^{(j)})^T x^{(i)}$. Za dani skup podataka koji se sastoji od ocjena nekih korisnika za neka piva, želimo naučiti vrijednosti vektora parametara $x^{(1)}, \dots, x^{(n_b)}, \theta^{(1)}, \dots, \theta^{(n_u)}$ tako da predviđanja najbolje odgovaraju stvarnim vrijednostima (želimo minimizirati kvadratnu grešku). U tu svrhu tražimo parametre koji minimiziraju sljedeću funkciju greške:

$$J(x^{(1)}, \dots, x^{(n_b)}, \theta^{(1)}, \dots, \theta^{(n_u)}) = \frac{1}{2} \sum_{(i,j): R(i,j)=1} ((\theta^{(j)})^T x^{(i)} - y^{(i,j)})^2. \quad (1)$$

Preporuka piva

Nakon što smo kolaborativnim filtriranjem naučili vrijednosti parametara, za željenog korisnika j predviđamo ocjenu $(\theta^{(j)})^T x^{(i)}$ za svako pivo i te od onih piva koja korisnik j nije ocijenio vraćamo listu piva s najvećim predviđenim ocjenama.

Opis funkcija

initialize(beers, users, ratings) Funkcija koja prima listu imena piva i korisnika te rječnik postojećih ocjena, a vraća inicijalizirane parametre spremjene redom u jednu listu te matrice Y i R .

normalize(Y, R) Funkcija koja normalizira matricu Y tako da svakom elementu i -tog retka matrice Y , ako je na istom mjestu u matrici R upisana vrijednost 1, oduzme prosječnu ocjenu piva i . Uz normaliziranu matricu Y , funkcija vraća i vektor prosječnih ocjena.

cost(params, Y, R, num_beers, num_users) Funkcija koja vraća vrijednost funkcije greške J . 1

predict(user, params, beers, users, ratings, mean) Funkcija koja korisniku s imenom "user" vraća listu neocjenjenih piva s najvećim predviđenim ocjenama, odnosno listu parova (predviđena ocjena, ime piva).