Recommender sistemi i njihov značaj u IT svetu

**Sadržaj –** Mnoge IT kompanije iz Silikonske doline danas pokušavaju da poboljšaju sisteme preporuka (eng. recommender systems). Razlog za unapređenje istih jeste njihov značaj u ukupnom prihodu tih kompanija. Ukoliko se uzme u obzir Amazon može se uvideti da što se bolje korisniku preporuči sadržaj za kupovinu on će više para potrošiti. Sistemi preporuka su tu da privuku korisnikovu pažnju tako što pronalaze šta je ono što njega zanima.

**Abstract** – Today, many iT companies from the silicon valley are looking for a better recommender systems. The main reason behind this is the significant role recommender systems have in revenue generation. If you take a look at the Amazon for example you can notice that depending on how well you recommend content to a user it will reflect on his will to spend money. So, recommender systems are there to recommend the best possible content to the user in that way bringing more money to a company.

**Uvod**

Danas se korisnici sreću sa sve većom količinom podataka na internetu. Ono što se pojavilo kao problem jeste kako omogućiti korisniku da na što lakši način dođe do sadržaja koji ga zanima. Na primer, sajt koji sadrži filmove ima za cilj da korisniku preporuči filmove koji bi se njemu svideli. Preporuka sadržaja korisnicima se uglavnom vrši bazirano na prethodnim iskustvima sa korisnicima.

**Algoritam**

Ovde će biti prikazan collaborative filtering algoritam koji ima tu osobinu da sam uči skup osobina koje su ključne za problem. Tokom dužeg vremenskog perioda se posmatra aktivnost korisnika i tako uči o korisnicima i njihovim interesovanjima. Algoritam će biti prikazan na primeru sajta koji sadrži filmove. Informacije koje su poznate su korisnici, i njihove ocene na neke filmove. Zadatak jeste predvideti koje ocene bi korisnici dali na ostale filmove. Za svakog korisnika treba naučiti parametar  (parametar korisnika). Predviđa se da je korisnik ocenio film sa , gde je vektor feature-a filma . Neka je funkcija ocena koje su korisnici dali  (ocena korisnika na film ). Ukoliko su poznati feature-i filma (npr. stepen komičnosti, stepen akcije u filmu) onda se zadatak svodi na nalaženje parametara korisnika. Parametri korisnika se dobijaju tako što se nekom numeričkom metodom povuče prava koja minimalno odstupa od  tačaka u prostoru (problem linearne regresije). Da bi se naučio (parametar za korisnika ) treba naći takvo koje dovodi do minimalne vrednosti sledećeg izraza:

ukoliko je korisnik j ocenio film i.Da bi se došlo do parametara (parametara za sve korisnike) potrebno je sumirati prethodan izraz po svim korisnicima:

Za nalaženje minimuma prikazane funkcije može se koristiti neka numerička metoda (npr. gradient descent). Ova metoda se svodi na pronalaženje najniže tačke (minimuma) funkcije tako što se iterativno (u koracima) putuje na “dole“ što kraćim putem. To se postiže tako što se kreće po pravcu tangente na krivu (izvoda funkcije u tački). Uspomoć opisane metode dolazi se do sledećeg skupa jednačina:

Na ovaj način polazeći od poznatih vektora feature-a filmova se dolazi do parametera korisnika. Problem se može postaviti i obrnuto, ukoliko su poznati parametri korisnika treba naći vektore feature-a filmova. Ovo je moguće na sličan način kao i već opisan. Moguće je iterativno učiti feature-e filmova, pa na osnovu bolje procene feature-a filmova sračunati opet parametre korisnika, i tako u krug, na taj način poboljšavajući procene. Ono što se uobičajeno dešava jeste da nisu poznate ni informacije o parametrima korisnika ni vektori feature-a filmova, već samo pojedine ocene na filmove. Collaborative filtering algoritam daje rešenje za ovu situaciju. Prvo se inicijalizuju parametri korisnika i vektori feature-a filmova na neke male random vrednosti. Zatim se minimizuje izraz:

po parametrima korisnika i po feature-ima filmova:

Na ovaj način omogućava se iterativno učenje parametra korisnika i feature-a filmova koji vremenom konvergiraju u sve bolje procene.

**Zaključak**

Kada se sračunaju parametri korisnika i bitne karakteristike raspoloživih filmova, moguće je uvideti koji filmovi su jako slični filmovima koje je korisnik ocenio dobro, na taj način se daje preporuka korisniku i u interesu je da ta preporuka bude što bolja. Dobrom preporukom postiže se da su korisnici zadovoljniji prilikom upotrebe programa, sajta.

**Reference**

[1] Machine learning kurs <https://www.coursera.org/course/ml>