

Predlog projekta

Članovi tima: Luka Nikolić, SW64/2017

Motivacija

Pronaći film koji nam se sviđa u današnje vreme sa izobiljem sadržaja je monumentalan zahvat. Potrebno je razviti aplikaciju koja će korisniku na osnovu njegovih preferencija predložiti najbolji film.

Opis problema:

Problem predlaganja sadržaja su probale da rese mnoge kompanije i platforme. Netflix je razvio kompleksan machine learning algoritam sa velikim brojem parametara podataka [1]. Sa druge strane, rad [2] uvodi rule based model koji predviđa preferencije grupe korisnika na osnovu njihovih istorija i ocena.

Predlažem sistem Movie recommender koji će, koristeći Drools rule engine pored korisničkih omiljenih žanrova i ocena, uzeti u obzir ključne reči koje dublje opisuju filmove od njihovih kategorija i ocena. Ključne reči su generalan pojam koji slobodnije i preciznije mogu da opišu atmosferu, iskustvo i osećanja dok gledamo film od drugih metrika. Iz tog razloga smatram da mogu da se koriste za kvalitetnije predviđanje filmova koji bi se korisniku svideli.

Metodologija rada

Ulazi:

- Omiljeni žanrovi – *Horror, Comedy, Drama...*
- Ključne reči – *based on comic, coming of age, superhero...*
- Recenzije – ostavljene od strane korisnika
 - Ocena – 1-10
 - Komentar – Opcioni tekst
- Korisnička istorija – kolekcija pregledanih filmova

Izlaz:

- Lista predloga sortirana prema stepenu poklapanja sa korisničkim preferencijama

Baza znanja:

Sistemu je potrebna kolekcija filmova koja se može prikupiti pomoću skrejpera sa online baza podataka filmova. Takođe su potrebne kolekcije ključnih reči i žanrova vezanih za

konkretne filmove, koje se mogu prikupiti na isti način. Rule engine će učitati sve činjenice prilikom pokretanja aplikacije i formiranja baze. Nakon bilo koje izmene postojećih činjenica od strane pravila, baza podataka aplikacije će takođe biti ažurirana

Primer rezonovanja

Primer glavnog skupa pravila (forward chaining)

Pravilo A1 – Filtriraju se filmovi na osnovu korisnikovih odabranih žanrova. Prilikom ovog koraka se generišu *MoviePrediction* činjenice koje se koriste u narednim koracima

Pravilo A2 – Za prethodno kreirane *MoviePrediction* činjenice, povećava se ranking prema prosečnoj oceni svih recenzija za film

Pravilo A3 – Za prethodno kreirane *MoviePrediction* činjenice, povećava se ranking prema stepenu poklapanja korisnikovih odabranih ključnih reči sa filmom

Pravilo A4– Ukoliko korisnik ima istoriju pregledanih filmova, povećava se ranking filmovima koji dele ključne reči sa filmovima iz korisnikove istorije (Opciono može da se množi sa ocenom koju je korisnik ostavio na tom filmu)

Primer međusobne isključivosti

Ukoliko korisnik nije naveo omiljene žanrove, izvršiće se sledeće pravilo:

Pravilo B1 – Uzimaju se nekoliko najpopularnijih žanrova na osnovu recenzija filmova. Zatim se kreiraju *MoviePrediction* činjenice i nastavlja se tok od pravila A2

Konkretni primer

Ulazi korisnika

- Omiljeni žanrovi: Horror, Mystery, Drama
- Ključne reči: suspenseful, slasher, serial killer
- Nema recenzija i nema istoriju gledanja

Baza znanja sadrži sledeće filmove: Silence of the Lambs (7.5), Inception(8.2), Airplane(7.9)

Izvršavanje pravila:

A1 - Filtriraju se filmovi iz baze znanja na osnovu izabranih žanrova. Rezultat ovog koraka su filmovi Silence of the Lambs i Inception

A2 – Na osnovu recenzija svih korisnika Inception dobija više poena (8.2)

A3 – Na osnovu ključnih reči suspenseful i serial killer Silence of the Lambs dobija 2 poena (ukupno 9.5), a Inception 1 zbog suspenseful (ukupno 9.2)

Pošto korisnik nema istoriju, ne izvršava se korak A4 i najbolji film je Silence of the Lambs zbog najviše bodova

Literatura:

1. Bennett, James, and Stan Lanning. "The netflix prize." Proceedings of KDD cup and workshop. Vol. 2007. 2007. Retrieved from:
<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.115.6998&rep=rep1&type=pdf>
2. Sharma, Mugdha, Laxmi Ahuja, and Vinay Kumar. "A novel rule based data mining approach towards movie recommender system." Journal of Information and Organizational Sciences 44.1 (2020): 157-170. Retrieved from:
<https://hrcak.srce.hr/file/348076>