

SBNZ Predlog Projekta

Music Recommendation System

Hajduk Dušan sw40-2018, Nikola Petrović sw34-2018

Motivacija

Kada smo u potrazi za novom muzikom koju nismo do sada čuli i pri tome želimo da ona ima slične karakteristike kao i muzika koju generalno slušamo, nikad ne znamo kojoj aplikaciji se obratiti. Ukoliko nam se sviđa neki izvođač ili album, postoji velika šansa da će nam se svideti još pesama od tog istog izvođača. Zbog toga potreban nam je sistem za preporuku muzike na koji možemo da se oslonimo.

Pregled Problema

[Jedno postojeće rešenje](#) - Naše rešenje će se razlikovati od navedenog po tome što ćemo pored ličnih informacija i ukusa za muziku nadgledanog korisnika uzimati u obzir i šta ostali korisnici slični našem korisniku slušaju.

Probleme koje ćemo rešiti:

- Na osnovu pravila generisati korisniku predloge pesama
- Na osnovu slušanja svih korisnika generisati liste najpopularnijih pesama

Metodologija Rada

- Ulaz
 - U slučaju preporuke za korisnika, ulaz je informacije o tom korisniku, ocene i neki dodatni upiti koje je uneo.
 - Informacije o ostalim korisnicima sistema
 - Za izveštaje ulaz generalna sve informacije o pesmama, muzicarima.
 - Trenutna rangiranja pesama, muzičara i žanrova
- Izlaz
 - Izlaz za korisnika su preporučene pesme na osnovu nadgledanog korisnika
 - Izlaz za admina su izveštaji za najpopularnije pesme, muzičare i žanrove
- Baza znanja
 - Korisnik - Osnovni podaci, istorija slušanja pesama, istorija ocena
 - Godine - 12-20 - Tinejdzer; 20-40 - Srednjih godina; 40-60 - Starijih godina; 60-100 Penzioneri

- Pisma - Dužina, izvodjač, listu žanrova, listu ocena, godina izlaska, broj slušanja, listu osobina (dinamicna, srecna, spora, brza...)
- Izvodjač - Lista pesama, lista žanrova
- Slusanje - Datum slusanja, pesma i korisnik
- Baza pesama, izvođača i žanrova

Pravila

1. Preporuka pesama prema korisnikovim podacima

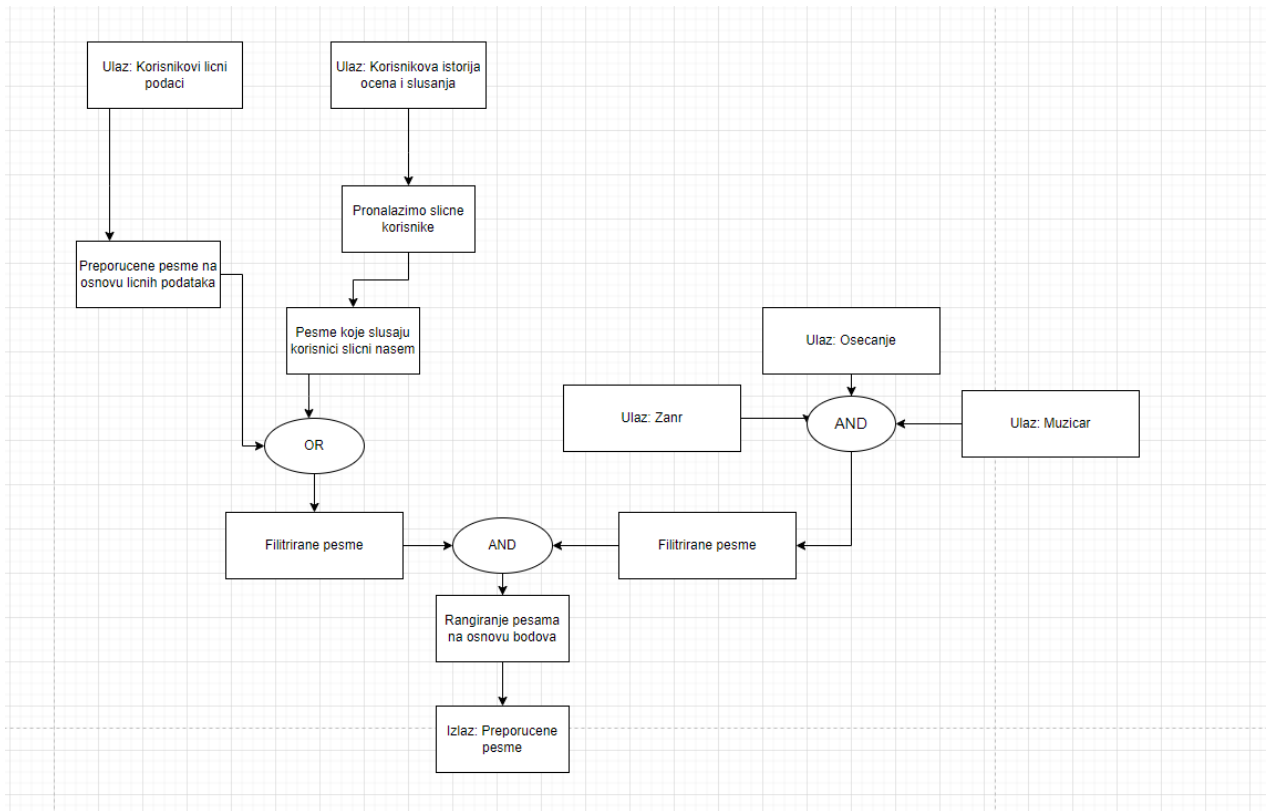
Korisniku se preporučuje 5 najbolje rangiranih pesama.

Primer preporuke:

Na ulazu uzimamo korisnikove podatke poput istorije slušanja i ocena. Korisnik može takodje na ulazu da unese neke dodatne informacije o tome kakav tip (srećna, brza..) pesama želi, o zanru ili od kog muzicarai.

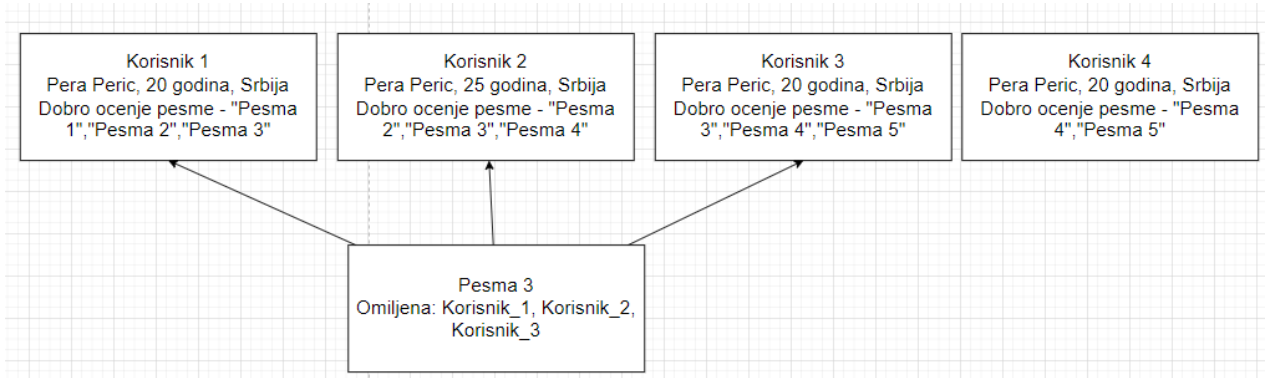
Vršimo 2 rangiranja:

- na osnovu korisnikovih podataka
 - unetih informacija
-
- Bodovanje po žanrovima
Povećanje bodovanja svim pesmama koje sadrže žanr koji je korisnik slušao u prošlosti.
Formula: **trenutni_bodovi na pocetku su 0; trenutni_bodovi + 1**
 - Bodovanje po muzicarima
Povećanje bodova pesmama u zavisnosti od odnosa slušanosti tog izvođača od strane korisnika i prosečne ocene tih pesama koje je slušao.
Formula: **trenutni_bodovi + (broj_odlušanih_pesama/100) * 0.2 * prosečna_ocena_pesama**
 - Bodovanje po vremenu trajanja
Povećanje bodovanja svake pesme prema trajanju pesama koje je korisnik slušao u prošlosti.
Formula: **trenutni_bodovi + 0.2* abs(trajanje_pesme - prosečno_trajanje_pesama_korisnika)**
 - Bodovanje po sličnom ukusu ostalih korisnika
Uzimamo najslusanije i dobro ocenje pesme korisnika i pronalazimo korisnike sa sličnim "ukusom". Uzimamo njihove najslusanije i dobro ocenje pesme i dodajemo bodove
Formula: **trenutni_bodovi + 0.1*broj_korisnika_kojima_je_ovo_najslusanija_pesma**
 - Rangiranje pesama
Na osnovu skupljenih bodova sortiramo pesme

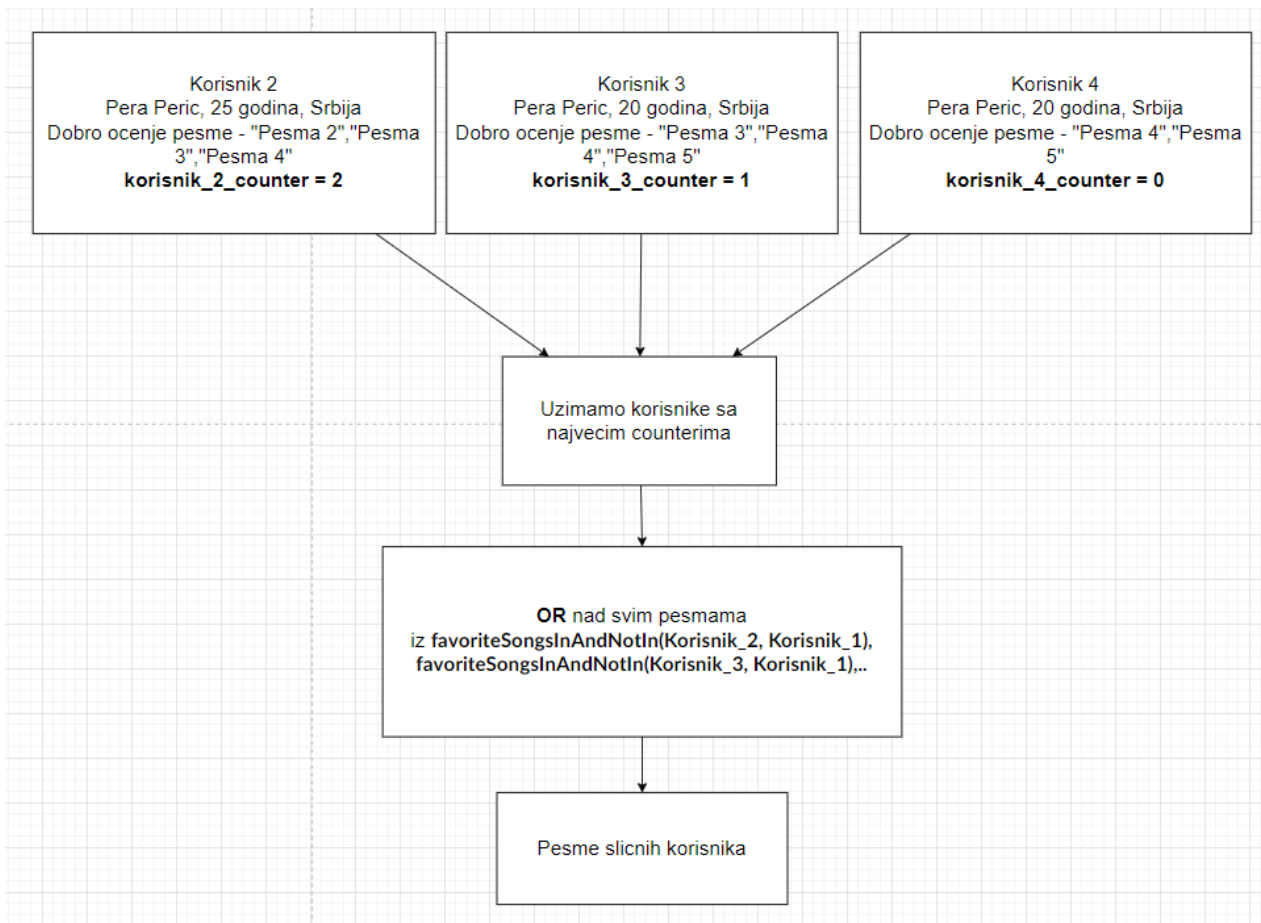


Primer pronalazenja slicnih korisnika:

- Ulaz : **Korisnik_1** (nadgledani korisnik) i njegove omiljene pesme **Pesma 1**, **Pesma 2** i **Pesma 3** i njegova istoriju slusanja



- Za svaku pesmu backward chaining trazimo korisnike koji imaju tu pesmu u omiljenim:
 $\text{songsContainedIn}(\text{Pesma 1}, \text{Korisnik_2}) \Rightarrow$ nema je u omiljenim preskacemo
 $\text{songsContainedIn}(\text{Pesma 2}, \text{Korisnik_2}) \Rightarrow$ ima je u omiljenim
 $\text{korisnik_2_counter} += 1, \dots$



- **korisnik_2_counter = 2** posto ima **Pesma 2** i **Pesma 3** u omiljenim
- Uzimamo korisnike sa najviše sličnih pesama našem korisniku
- Koristimo **favoriteSongsInAndNotIn(Korisnik_2, Korisnik_1)** pronadjemo pesme koje su omiljene **Korisnik_2**, a **Korisnik_1** nije slusao
- Izlaz: Pesme koje su omiljene sličnim korisnicima, a naš korisnik ih nije slusao

Jedan primer:

Uneli smo žanr i osećanje i uzimamo godine korisnika

- 19 godina => tinejdžer
- Tinejdžer => Novije pesme
- Novije pesme => Pesme u zadnjih 5 godina
- Izabran žanr => Pesme izabranog žanra
- Izabrano osećanje => Pesme izabranog osećanja

U svakom koraku prenesemo dobijeni rezultat dalje

Dobijamo filtrirane pesme.

Drugi primer:

Korisnik nije uneo dodatna polja

Uzimamo godine korisnika 25 => Srednjih godina

Srednjih godina => Pesme sa osobinom za slusaoce srednjih godina

Uzimamo slusanja i ocena => Filtriramo najslusanije i najbolje ocenjene pesme

Uzimamo ostale korisnike => Na osnovu izdvojenih pesama i licnih podataka pronadjemo slicne korisnike

Uzmemo najslusanije druge pesme drugih korisnika => Spojimo rezultate sa gore dobijenim za licne podatke

Rangiramo pesme na osnovu bodova => Rangirane pesme

2. Pregled najpopularnijih pesama

Korisnik ima uvid u najpopularnije pesme, što se određuje na sledeći način:

- a. Filter za pronalaženje najpopularnije u odredjenom vremenskom periodu. Filter se prosledjuje pri daljim evaluacijama.
- b. Klasifikujemo pesme na osnovu dužine.
Podela je:
 - 1) Pesme do 2 minuta
 - 2) Pesme od 2-5 minuta
 - 3) Pesme preko 5 minuta
- c. Pri rangiranju pesama na osnovu podele pod **b.** pesma pod kategorijom **1)** dodajemo **-0.7** bodova pesmama pod kategorijom **2)** dodajemo **+0.9** bodova, a pesmama pod kategorijom **3)** dodajemo **+0.5** bodova.

Formula: $10 + \text{bodovi_na_osnovu_duzine}$

- d. Pri rangiranju pesama računamo prosečnu ocenu svih pesama. Na osnovu prosečne ocene svih pesama svakoj pesmi koja je 10% iznad proseka dajemo ***1.2** bodova, pesmama koje su u opseg [-10%,+10%] od proseka ne menjamo bodove, a pesmama koje su -10% od proseka dajemo ***0.8** bodova.

Formula: $(10 + \text{bodovi_na_osnovu_duzine}) * \text{bodovi_na_osnovu_pr_ocene}$

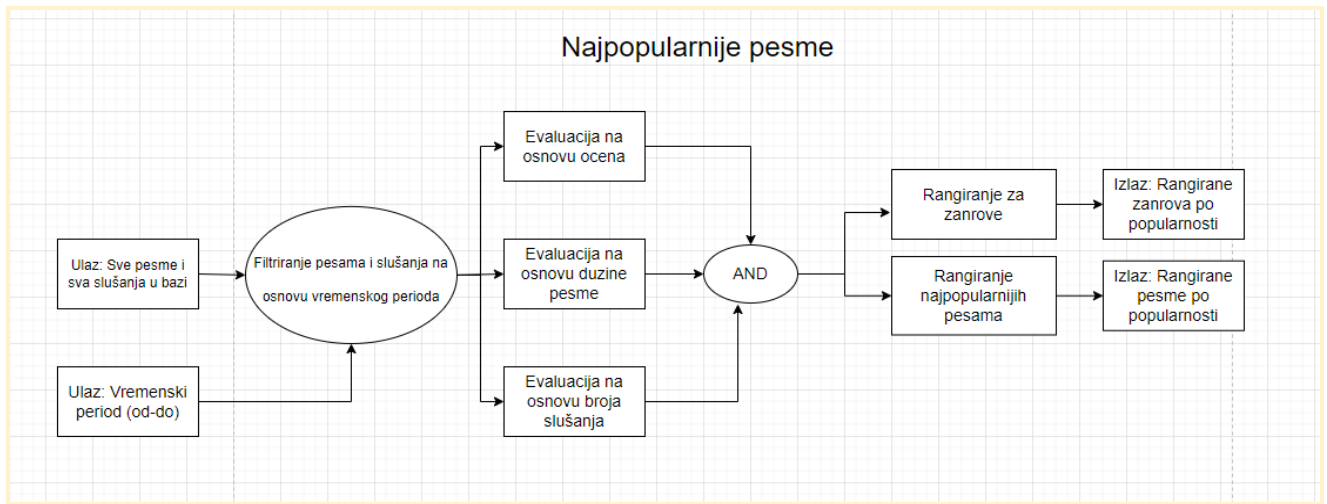
- e. Pri rangiranju pesama uzimamo broj slušanja svih pesama. Na osnovu prosečne broja slušanja svih pesama svakoj pesmi koja je 10% iznad proseka dajemo **+0.3** bodova, pesmama koje su u opseg [-10%,+10%] od proseka ne menjamo bodove, a pesmama koje su -10% od proseka dajemo **-0.3** bodova.

Formula: $(10 + \text{bodovi_na_osnovu_duzine}) * \text{bodovi_na_osnovu_pr_ocene} + \text{bodovi_na_osnovu_broja_slusanja}$

Na osnovu napisane formule možemo generalno rangirati pesme.

- f. Pomoću izračunatih bodova pesama možemo takodje rangirati najuticajnije žanrove. Žanrovi sa manje od 50 pesama skaliramo sa 0.8

Formula: $\text{zbir_pr_ocena_pesama_žanra} / \text{broj_pesama_žanra} * \text{skaliranje_na_osnovu_br_pesama}$



CEP

Svaki korisnik ima pravo samo jednu ocenu za svaku pesmu da napravi. Ukoliko je u prošlosti ocenio pesmu može promeniti svoju ocenu jednom u svaka 2 minuta. Ovim mehanizmom sprečavamo učestalo menjanje ocene od strane korisnika.

Rangiranje pesama, muzičara i žanrova ukoliko imamo veliki broj podataka proces može biti zahtevan i spor za obradu. Da ne bismo svaki put kad se odsluša ili oceni neka nova pesma ponovo računali sva rangiranja na nekom odredjenom intervalu ćemo izvršavati računanje novih rangiranja.

Literatura

https://www.researchgate.net/publication/220894443_COMUS_Ontological_and_Rule-Based_Reasoning_for_Music_Recommendation_System

<https://github.com/shivamawasthi14/Music-Recommendation-System>