

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 sa tog albuma ili 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.

Problemi 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, pesme koje su ostali korisnici dobro ocenili, a slušali su pesme kao i nadgledani korisnik.
 - Za izveštaje ulaz generalna sve informacije o pesmama, muzicarima.
 - Trenutna rangiranja pesama, muzičara i žanrova
- Izlaz
 - Izlaz za korisnike 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
 - Pesa - Dužina, izvodjač, listu žanrova, listu ocena, godina izlaska, broj slušanja
 - Izvodjač - Lista pesama, lista žanrova

- Baza pesama, izvođača i žanrova

Pravila

1. Preporuka pesama prema korisniku

Korisniku se preporučuje 5 najboljih pesama rangiranih na osnovu sledećeg bodovanja:

- Bodovanje po žanrovima
Povećanje bodovanja svim pesmama koje sadrže žanr koji je korisnik slušao u prošlosti.
Formula: $\text{trenutni_bodovi} + 1$
- Bodovanje po autorima
Povećanje bodova pesmama u zavisnosti od odnosa slušanosti tog izvođača od strane korisnika i prosečne ocene tih pesama koje je slušao.
Formula: $\text{trenutni_bodovi} + (\text{broj_odlušanih_pesama}/100) * 0.2 * \text{prosečna_ocena_pesama}$
- Bodovanje po popularnosti
Svaka pesma se boduje prema korisnicima koji su slušali istu pesmu u prošlosti
Formula: $\text{trenutni_bodovi} + 0.5 * (\text{broj_jedinstvenih_pregleda} / \text{broj_korisnika})$
- Bodovanje po vremenu trajanja
Povećanje bodovanja svake pesme prema trajanju pesama koje je korisnik slušao u prošlosti.
 $\text{razlika_u_trajanju} = \text{abs}(\text{trajanje_pesme} - \text{prosečno_trajanje_pesama_korisnika})$
Formula: $\text{trenutni_bodovi} + 0.2 * (1 - \text{razlika_u_trajanju}) \text{ (while } \text{razlika_u_trajanju} < 1)$
- Bodovanje po sličnom ukusu ostalih korisnika
Svaka pesma se boduje prema odnosu istih slušanih pesama trenutnog korisnika i svih ostalih korisnika sistema koji su slušali tu pesmu kao i njihove ocena te pesme.
Formula: $\text{trenutni_bodovi} + (\text{broj_istih_pesama}/100) * (\text{ocena_tog_korisnika}/10)$

2. Pregled najuticajnijih pesama

Korisniku ima uvid u najuticajnije pesme, što se određuje na sledeći način:

- a. Klasifikujemo pesme na osnovu dužine.

Podela je:

- 1) Pesme do 2 minuta
- 2) Pesme od 2-5 minuta

3) Pesme preko 5 minuta

- b. Pri rangiranju pesama na osnovu podele pod a. 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: **$10 + \text{bodovi_na_osnovu_duzine}$**
- c. 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: **$(10 + \text{bodovi_na_osnovu_duzine}) * \text{bodovi_na_osnovu_pr_ocene}$**
- d. 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: **$(10 + \text{bodovi_na_osnovu_duzine}) * \text{bodovi_na_osnovu_pr_ocene} + \text{bodovi_na_osnovu_broja_slušanja}$**
Na osnovu napisane formule možemo generalno rangirati pesme.
- e. Pomoću izračunatih bodova pesama možemo takodje rangirati najuticajnije muzičare na osnovu prosečne ocene svih njihovih pesama. Muzičari koji imaju više od 20 pesama bodove množimo sa 1.3, a muzičare koji imaju manje sa 0.8.
Formula: **$\text{zbir_pr_ocena_pesama_muzičara} / \text{broj_izdatih_pesama} * \text{skaliranje_na_osnovu_br_pesama}$**
- 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.