**Stvaranje i usporedba prediktivnih modela preporuke filma iz MovieLens dataseta.**

Projektna dokumentacija

Verzija 2.0

Sadržaj

1. Puni naziv projekta 4

2. Skraćeni naziv projekta 4

3. Opis problema/teme projekta 4

4. Cilj projekta 4

5. Voditelj studentskog tima 4

6. Rezultat(i) 4

7. Slični projekti 4

8. Resursi 4

9. Glavni rizici 5

10. Smanjivanje rizika 5

11. Glavne faze projekta 5

12. Struktura raspodijeljenog posla (engl. *Work Breakdown Structure* - WBS) 5

13. Kontrolne točke projekta 5

14. Gantogram 5

15. Zapisnici sastanaka 5

Puni naziv projekta

Stvaranje i usporedba prediktivnih modela preporuke filma iz MovieLens dataseta.

# Skraćeni naziv projekta

Prediktivni model preporuke filma

# Opis problema/teme projekta

Eksploratorna analiza te razvijanje raznih prediktivnih modela preporuke filma algoritmima strojnog učenja pomoću MovieLens dataseta kao izvor filmova i ratinga. Modeli bi preporučivali nove filmove korisniku s obzirom na podatke koje ima o njemu kao što su koje žanrove preferira, koje filmove je dobro ocijenio, a koje loše te s obzirom na podatke vrlo sličnih korisnika. Razvijene prediktivne modele bi testirali, ispitali te ih međusobno usporedili te iznijeli zaključke.

# Cilj projekta

Razvijanje i usporedba prediktivnih modela koji će preporučiti filmove korisniku na osnovi podataka koje ima o njemu i ostalim korisnicima sličnog ukusa.

Predviđa se da projekt traje 3 mjeseca.

# Voditelj studentskog tima

Matej Čubek

# Rezultat

Rezultat projekta su prediktivni modeli koji za bilo kojeg korisnika iz baze podataka mogu preporučiti novi film koji bi mu se s velikom vjerojatnošću svidio te njihova međusobna usporedba.

# Slični projekti

<https://github.com/AshwathSalimath/Movie-ratings-prediction-using-MovieLens-Dataset>

<https://github.com/AIGamer28100/Movie-Lens-Dataset-Visualisation-and-Prediction>

<https://datascience-enthusiast.com/Python/cs110_lab2_als_prediction.html>

# Resursi

Tablica ljudskih resursa

|  |  |  |  |
| --- | --- | --- | --- |
| **Ime i prezime** | **E-mail adresa** | **GSM broj** | **Napomene** |
| Iva Bokšić | Iva.boksic@fer.hr | +385976969143 | Poznavanje R |
| Tomislav Grbeša | Tomislav.grbesa@fer.hr | +385977259354 | Predznanje datascienca |
| Matej Čubek | Matej.cubek@fer.hr | +385992066001 | Znanje Jave i pythona |

# Glavni rizici

* Nedostatak vremena za upoznavanje s znanosti o podacima i izradom prediktivnog modela.
  + Posljedica rizika je nedovoljno dorađen prediktivni model.
* Premalo prijašnjeg znanja za izradu prediktivnog modela
  + Posljedica rizika je nedovoljno dorađen prediktivni model.
* Nedostatak raspoloživih materijala za učenje
  + Nemoguće odraditi projektni zadatak

# Smanjivanje rizika

* Bolja organizacija vremena za projekt.
* Učenje novih materijala kako bi dopunili znanje.
* Potraga za novim materijalima na internetu ili traženje pomoći od mentora
* Česti sastanci radi podjele posla i rasprave što je i kako do sada napravljeno

# Glavne faze projekta

Projekt je podijeljen na vremenske faze od kojih svaka traje dva do tri tjedna.

* 1. Faza: Stvaranje tima, organizacija i odabir tehnologija.
  2. Faza: Odabir dataseta i njegova eksploracija i analiza.
  3. Faza: Obrada podataka i priprema za algoritme strojnog učenja.
  4. Faza: Izrada modela.
  5. Faza: Usporedba i testiranje modela.
  6. Faza: Objava rezultata.

# Struktura raspodijeljenog posla (engl. *Work Breakdown Structure* - WBS)

# Kontrolne točke projekta (engl. *milestones*)

**Tablica kontrolnih točki projekta**

|  |  |  |  |
| --- | --- | --- | --- |
| **Kontrolne točke** | **Planirani datum** | **Realizirani datum** | **Status projekta** |
| Predaja prve verzije plana projekta | 6.11.2020. | 6.11.2020. | Određen cilj projekta i napravljen plan. |
| Gotova analiza dataseta | 1.12.2020. | 2.12.2020. | Eksploratorna analiza podatka izvršena. |
| Razvijeni prediktivni modeli | 6.01.2021. | 8.01.2021. | Modeli napravljeni. |
| Uspoređeni prediktivni modeli | 12.01.2021. | 14.01.2021 | Modeli uspoređeni i napravljena analiza. |
| Objava rezultata i predaja | 17.01.2021. | 21.01.2021. | Projekt završen. |

# Gantogram

# Zapisnici sastanaka

* 1. Sastanak : 22.10.2020. 19h, Slack
     + Sudjelovali: Iva Bokšić, Tomislav Grbeša, Matej Čubek
     + Upoznavanje tima i rasprava o trenutnim znanjima pojedinog člana
     + Razmjena korisnih materijala za upoznavanje s projektom
  2. Sastanak : 29.10.2020. 18h, Slack,
     + Sudjelovali: Iva Bokšić, Tomislav Grbeša, Matej Čubek
     + Podjela rada pri izradi dokumentacije
  3. Sastanak : 01.12.2021. 09h, Zoom,
     + Sudjelovali: Iva Bokšić, Tomislav Grbeša, Matej Čubek
     + Početci eksploratorne analize i osnovni model
  4. Sastanak : 11.12.2020. 10:30h, Teams,
     + Sudjelovali: Iva Bokšić, Tomislav Grbeša, Matej Čubek, Mentor
     + Kontrolna točka
  5. Sastanak : 14.12.2020. 18h, Zoom,
     + Sudjelovali: Iva Bokšić, Tomislav Grbeša, Matej Čubek
     + Brainstorming vrsta prediktivnih modela i intrgeacije
  6. Sastanak : 29.10.2020. 12, Slack,
     + Sudjelovali: Iva Bokšić, Tomislav Grbeša, Matej Čubek
     + Potres
  7. Sastanak : 05.01.2021. 10h, Teams,
     + Sudjelovali: Iva Bokšić, Tomislav Grbeša, Matej Čubek
     + Izgradnja modela
  8. Sastanak : 08.01.2021. 11h, Teams,
     + Sudjelovali: Iva Bokšić, Tomislav Grbeša, Matej Čubek
     + Pregled kroz modele
     + Testiranje modela
  9. Sastanak : 13.01.2021. 10h, Teams,
     + Sudjelovali: Iva Bokšić, Tomislav Grbeša, Matej Čubek
     + Dokumentiranje
  10. Sastanak : 16.01.2021. 10h, Teams,
      + Sudjelovali: Iva Bokšić, Tomislav Grbeša, Matej Čubek
      + Dodavanje markdowna
  11. Sastanak : 17.01.2021. 10h, Teams,
      + Sudjelovali: Iva Bokšić, Tomislav Grbeša, Matej Čubek
      + Dogovor o završnim točkama
  12. Sastanak : 20.01.2021. 10h, Teams,
      + Sudjelovali: Iva Bokšić, Tomislav Grbeša, Matej Čubek
      + Finalno uljepšavanje
  13. Sastanak : 21.01.2021. 10h, Teams,
      + Sudjelovali: Iva Bokšić, Tomislav Grbeša, Matej Čubek
      + Proba prezentiranja

|  |  |
| --- | --- |
| **Suglasan s dokumentom (potpisuju članovi tima):** |  |
| Iva Bokšić | Datum: \_\_\_\_\_\_\_\_\_\_\_\_ Potpis: \_\_\_\_\_\_\_\_\_\_\_\_ |
| Tomislav Grbeša | Datum: \_\_\_\_\_\_\_\_\_\_\_\_ Potpis: \_\_\_\_\_\_\_\_\_\_\_\_ |
| Matej Čubek | Datum: \_\_\_\_\_\_\_\_\_\_\_\_ Potpis: \_\_\_\_\_\_\_\_\_\_\_\_ |

|  |  |
| --- | --- |
| **Odobrio(potpisuje nastavnik):**  izv. prof. dr. sc. Damir Pintar |  |

|  |  |
| --- | --- |
| Datum: \_\_\_\_\_\_\_\_\_\_\_\_ | Potpis: \_\_\_\_\_\_\_\_\_\_\_\_ |