

مبانی بازیابی اطلاعات و جستجوی وب

نیم سال دوم ۱۴۰۱-۱۴۰۲ - کد ۲

پروژه پایانی

هدف از این پروژه، آشنایی با نحوه عملکرد سیستم‌های توصیه‌گر مبتنی بر دو روش پالایش همکارانه و مبتنی بر محتوا است. به این منظور شما می‌بایست بر اساس مجموعه داده‌ای که در اختیارتان قرار گرفته است، سیستمی طراحی نمایید که با دریافت نام یک فیلم، لیستی از ۱۰ فیلم مرتبط را نمایش دهد.

مجموعه داده:

این دیتاست شامل اطلاعات نظرات ۷۳۵۱۶ کاربر بر ۱۲۲۹۴ انیمه می‌باشد. هر کاربر با اضافه نمودن انیمه به لیست مشاهده‌شد های خود در وبسایت myanimelist.net می‌تواند یک امتیاز برای آن ثبت نماید که این دیتاست شامل همین مجموعه امتیازات می‌باشد.

در فایل فشرده‌ای که در اختیار دارید دو فایل با نام های anime.csv و rating.csv موجود است که در ادامه به معرفی آنها می‌پردازیم:

anime.csv

- anime_id: شناسه اختصاصی انیمه در وبسایت
- name: نام کامل انیمه
- genre: ژانر انیمه
- type: نوع انیمه
- episodes: تعداد اپیزودهای انیمه
- rating: میانگین امتیازات ثبت شده از ۱۰ برای انیمه
- members: تعداد اعضای جامعه‌ای که در گروه مختص به این انیمه عضو هستند.

rating.csv

- user_id: شناسه یکتای کاربر
- anime_id: شناسه یکتای انیمه
- rating: امتیاز ثبت شده کاربر برای انیمه

امتیازات ثبت شده برای هر انیمه از ۱۰ می‌باشد. مقدار ۱- به این معنا می‌باشد که کاربر انیمه را مشاهده، اما برای آن امتیازی ثبت ننموده است.

سیستم مبتنی بر پالایش همکارانه:

در این بخش شما می‌بایست با اعمال الگوی پالایش همکارانه بر امتیازاتی که کاربران برای فیلم‌ها ثبت نموده‌اند، به کاربر فیلم‌های جدیدی را پیشنهاد نمایید.

سیستم مبتنی بر روش مبتنی بر محتوا:

در این بخش شما می‌بایست با استفاده از فیلم‌هایی که کاربر مشاهده نموده و اطلاعات در دسترس از فیلم‌ها، مجموعه فیلم‌های مرتبط با آنها را به وی پیشنهاد نمایید.

در نظر داشته باشید که برای استفاده از داده‌های دیتاست، نیاز به پیش‌پردازش‌های مختلفی دارید که تشخیص پیش‌پردازش‌های لازم و اجرای آنها به عهده خودتان است.

خروجی:

در نهایت لازم است شما کد خود که می‌تواند به زبان دلخواه شما نوشته شده باشد را به همراه گزارشی که شامل توضیحات در مورد هر یک از روش‌های پیاده‌سازی شده و نمایش نتیجه به ازای ورودی‌های مختلف می‌باشد را در سامانه vu آپلود نمایید.

در صورتی که از زبانی غیر از پایتون استفاده می‌نمایید، نحوه اجرای کد پیاده‌سازی شده را در یک فایل توضیح دهید.