

سوال دوم)

در این بخش، سیستم‌های توصیه‌گر (recommender systems) را بر اساس امتیازات کاربران به کتاب‌های آمازون بررسی می‌کنیم. در ابتدا دیتاست از [اینجا](#) دانلود کنید. این دیتاست شامل ۴ ستون است:

- آیدی کاربر
- آیدی کتاب
- امتیاز
- زمان ثبت امتیاز

الف) در ابتدا بخشی از دیتاست به عنوان داده تست جدا کرده و در ادامه برای باقی دیتاست، بر اساس یکی از روش‌های Collaborative یا Latent factor based الگوریتم خود را طراحی و پیاده‌سازی کنید. تمام جزئیات در نظر گرفته‌شده و مراحل طراحی را گزارش دهید. برای نوشتن بخش LSH الگوریتم می‌توانید از توابع آماده استفاده کنید ولی باقی قسمت‌ها نیاز به کدنویسی از پایه دارند و در صورتی که نیاز به بهینه کردن تابع هزینه در الگوریتم خود دارید استفاده از کتابخانه‌ی MLlib مجاز نمی‌باشد.

ب) پس از طراحی سیستم توصیه‌گر، خطای RMSE برای بخش تست بدست آورید. می‌توانید جهت بهبود عملکرد قسمت (ج)، الگوریتم خود را به نحوی تغییر دهید که به کاهش این خطا منجر شود. (الگوریتم اولیه به همراه نتایج و مراحل تغییر منجر به کاهش این خطا گزارش کنید).

ج) در این بخش باید فایل Evaluation_P2 که در پوشه MDA_P2 قرار دارد را تکمیل کنید. کافی است برای هر آیدی یوسر/کتاب داده شده، امتیاز تخمین زده شده بجای "?" وارد کنید. (بخشی از نمره این تمرین به نتایج این بخش اختصاص دارد).