

Design Document

استفاده از الگوریتم K-Means برای خوشه بندی کاربران در مدل ریکامند به دلیل اینکه برای ساختن پیشنهاد ها برای هر کاربر اطلاعات تمام کاربران را از دیتابیس دریافت نکنیم. K-Means هر ۲۴ ساعت یکبار بر اساس دیتاهای جدید کاربران را دسته بندی میکند

برای ریکامند کردن از روش collaborative filter استفاده کردم که با استفاده از الگوریتم cosine similarity بردار های شباهت رو رسم کردم. دلیل انتخابم رایج بودن این الگوریتم و دقت بالاش برای اینکار بود. برای کاربرانی که دیتای کافی برای ساختن پیشنهاد ندارن محصولات پرتعداد بر اساس امتیازشون فیلتر و بهشون نشون داده میشه. اینجا بهتر بود از روش content based استفاده کنم اما متأسفانه زمان کم اوردم برای یاد گرفتنش.

محصولات پیشنهاد شده برای کاربران رو هم توی ردیس ذخیره میکنم که تا ۳ ساعت کش میشه و همچنین برای ساختن دیتا فریم ها از تاریخچه کاربرانی که داخل یک خوشه هستن استفاده میکنم.