

PH #5: Recommender System (Extra point)

سیستم پیشنهاد دهنده

در این بخش ما از شما انتظار داریم که با توجه به مفاهیم صحبت شده در قسمت Advanced AI و بحث های مرتبط با سیستم های پیشنهاد دهنده، با توجه به داده های موجود، یک سیستم پیشنهاد دهنده ارائه دهید. به طور کلی شما باید از دو روش فاکتوری سازی ماتریس یا Matrix Factorization و شبکه های عصبی برای اینکار استفاده کنید و نتایج این دو روش را با هم مقایسه کنید.

به عنوان یک نمونه از توضیح و کد برای روش فاکتورسازی ماتریس میتوانید از لینک زیر استفاده کنید:

[Matrix Factorization: A Simple Tutorial and Implementation in Python](#)

داده¹

داده ها، شامل دویست هزار (200.000) خرید توسط مشتریان است که همانند سایر داده های دیجیتال بی نام شده اند تا حریم شخصی مشتریان حفظ شود. این داده ها دارای زمان و موقعیت مکانی می باشد. در جدول زیر ستون های (ویژگی های) موجود در مجموعه داده را مشاهده می کنید.

مفهوم	ستونها
شناسه سفارش	ID_Order
شناسه مشتری	ID_Customer
شناسه محصول	ID_Item
تاریخ خرید	DateTime_CartFinalize
مقدار سود ناخالص (فروش)	Amount_Gross_Order
نام شهر خریدار	city_name_fa
تعداد سفارش	Quantity_item

پیش پردازش داده

برای شروع کار شما نیاز به یکسری پیش پردازش ها دارید، به عنوان مثال شناسه مشتری و محصولات در یک بازه مشخص نیست، که برای اینکار شما نیاز به اندیس گذاری آنها دارید. به طور مشخص، مثلا اگر شما در داده 10 کاربر دارید بهتر است که آنها را از 0 تا 9 اندیس دهید.

همچنین برای آموزش و ارزیابی شبکه نیاز به مجموعه داده آموزشی و تست دارید، برای این منظور شما باید ابتدا داده ها را براساس زمان خرید مرتب کنید (از قدیمی ترین به جدیدترین) و 150.000 رکورد ابتدایی را به عنوان مجموعه آموزشی و 50.000 رکورد انتهایی را به عنوان مجموعه تست قرار دهید.

¹ لطفا داده ها را تحت هیچ شرایطی منتشر نکنید.



ارزیابی

برای ارزیابی سیستم از دو معیار دقت² و فراخوانی³ استفاده کنید. بدین منظور باید نموداری برای مقایسه نتایج دو روش ارائه دهید که در این مقایسه و ارزیابی شما باید مقدار N در Top-N را 5، 10، 20 در نظر بگیرید. مقدار N در اینجا مشخص کننده تعداد آتم های پیشنهادی به کاربر است.

گزارش

هنگام ارسال کدهای برنامه، شما باید یک گزارش که هر بخش را به طور دقیق شرح داده باشد ارسال نمایید. به طور کلی پیشنهاد می شود از Jupyter Notebook استفاده کنید تا همزمان گزارش ها و نمودارهای مرتبط را چاپ کنید.

سوالات

در صورت هرگونه سوال لطفا با ایمیل من مکاتبه نمایید.

² Precision

³ Recall