گزارش مقاله جدید

به نام خدا

توضیح روش :

دیتاست : دیتاست شفارشات دیجی کالا

ابتدا روی داده rfm رو حساب میکنیم و بعد از اون outlier ها رو حذف میکنیم بعد در این مرحله هدف این روش پیشنهاد دادن محصولات مناسب به افرادی که از ماخرید داشته اند است.

به همین دلیل تنها افرادی را انتخاب میکنیم که تکرار خرید ان ها بیشتر مساوی 2 باشد(Frequency). که تعداد کاربران نهایی این برابر 9928 می باشد.

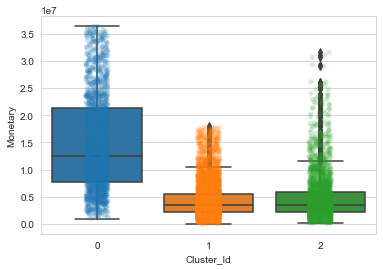
مرحله بعدی : با توجه به متفاوت بودن مقیاس، باید داده ها استاندارد سازی شود

مرحله ایجاد مدل : که با استفاده از مدل k-means ابتدا داده ها به خوشه ها مختلف تقسیم بندی میکنیم که تعداد خوشه بهینه برابر3 در نظر گرفته شده است. تعداد عضوهای خوشه ها برابر جدول زیر است

|  |  |
| --- | --- |
| شماره خوشه | تعداد عضو |
| 0 | 1689 |
| 1 | 5637 |
| 2 | 2602 |

ویژگی های هر خوشه :

خوشه صفر از نظر پولی (Monetary) از سایر خوشه هابیشتر است به این معنی که افراد این خوشه میانگین خرید های بالاتری نسبت به سایر خوشه ها داشته اند.



نمودار ویژگی های مربوط به R,F رو هم بعدا اضافه میکنم.

حالا با توجه به اعضای هر خوشه ، سفارشات هر خوشه جدا گانه بررسی میکنیم .

و جدولی تشکیل میدهیم که تعداد سطر های آن برابر با تعداد اعضا و تعداد ستون های آن برابر با تعداد محصولات .به عنوان مثال برای خوشه اول جدول برابر 1689\*10665 می باشد و برای مقادیری که جدول در صورتی که کاربر خرید داشته باشد مقدار جدول برابر، تعداد کالای خریداری شده، ودر غیر این صورت برابر با صفر می باشد.

سه جدول همانند جدول فوق تشکیل می شود (چون سه تا خوشه داریم )

حال در صورتی که کاربر بخواهند کالای با کد 5289 را خریداری کند با توجه به اینکه کاربر در کدام خوشه قرار دارد

پیشنهادات متفاوتی به او داده می شود.

به عنوان مثال کاربری در خوشه صفر خواهان خرید کالایی با کد 5289 می باشد که در این صورت با استفاده از ضریب همستگی ، محاسبه می شود با کدام کالاها بیشترین همبستگی را دارند

پیشنهادات کالای برای خوشه صفر : کالای 234341

پیشنهادات کالا برای خوشه یک :

پیشنهادات برای خوشه دو :

توضیحات بیشتر رو بعدا میذارم دوباره