

پروژه درس داده کاوی ترم دوم سال تحصیلی ۱۴۰۰ - ۱۳۹۹

سجاد ابراهیمی ۹۷۱۲۷۶۲۴۶۵

فاز دوم پروژه: شناسایی الگوهای مکرر

برای یک مجموعه داده که مربوط به دیتاست دیوار و دیجی کالا می باشد. موارد ذیل را برای این دیتاست ها انجام دهید.

در این فاز چهار گام کلی زیر باید صورت بپذیرند:

- **استخراج الگوهای قرار دادن پست فروش محصولات در شهرهای مختلف در دیتاست دیوار**

در این گام ما دسته بندی های هر محصول را همانند ایتm های سبد خرید شخص فرض کرده و سپس برای هر شهر به استخراج الگوهای مکرر میپردازیم. میزان `min_sup` را ۰/۲ در نظر میگیریم تا رکورد های شهر های مختلف نمایش داده شوند. به دلیل عدم پاسخ دهی `colab` کد را تنها بر روی ۲۵۰ رکورد از دیتاست اجرا میکنیم.

- **استخراج الگوهای فروش محصولات مختلف در دیتاست دیجی کالا**

در این گام ما سبد خرید مشتریان را ایجاد و ایتm ها را درون آن قرار داده و سپس به استخراج الگوهای مکرر میپردازیم. میزان `min_sup` را ۰/۰۰۰۵ در نظر میگیریم. به دلیل عدم پاسخ دهی `colab` کد را تنها بر روی ۲۰۰۰۰ رکورد از دیتاست اجرا میکنیم.

- **استخراج الگوهای مکرر میان روز انتشار پست و پلتفرم استفاده شده**

در این گام روز انتشار پست و پلتفرم را به عنوان آیتm های یک سبد در نظر گرفته و به استخراج الگوهای مکرر میپردازیم. میزان `min_sup` را ۰/۲ در نظر میگیریم. به دلیل عدم پاسخ دهی `colab` کد را تنها بر روی ۲۵۰ رکورد از دیتاست اجرا میکنیم.

- **استخراج الگوهای مکرر برای محصول و پلتفرم استفاده شده**

در این گام دسته بندی محصولات و پلتفرم را به عنوان آیتm های یک سبد در نظر گرفته و به استخراج الگوهای مکرر میپردازیم. میزان `min_sup` را ۰/۲ در نظر میگیریم. به دلیل عدم پاسخ دهی `colab` کد را تنها بر روی ۲۵۰ رکورد از دیتاست اجرا میکنیم.