

کشف الگوهای پرتکرار و قوانین انجمنی

توضیح تمرین:

روی دیتاست خود مراحل زیر را دنبال کنید:

۱. پیش‌پردازش داده‌ها:

- داده‌ها را بررسی و پاک‌سازی کنید.
- مقادیر گمشده (Missing Values) را شناسایی کرده و با روش مناسب آن‌ها را اصلاح کنید.
- داده‌ها را به فرم مناسب برای الگوریتم‌های کشف الگوهای پرتکرار تبدیل کنید (مقادیر ویژگی‌ها باینری شوند).

۲. کشف الگوهای پرتکرار:

- از الگوریتم Apriori برای کشف آیتم‌ست‌های پرتکرار استفاده کنید.
- الگوریتم FP-Growth را نیز پیاده‌سازی یا اجرا کنید و نتایج آن را با Apriori مقایسه کنید.

۳. استخراج قوانین انجمنی:

- با استفاده از آیتم‌ست‌های پرتکرار، قوانین انجمنی استخراج کنید.
- معیارهای ارزیابی مانند Support، Confidence و Lift را برای قوانین به‌دست‌آمده محاسبه و تحلیل کنید.

۴. تحلیل نهایی:

- نتایج را تحلیل و تفسیر کنید.
- به این پرسش پاسخ دهید: چه الگوهایی در این داده‌ها پرتکرار هستند و چه قوانینی از آن‌ها می‌توان استخراج کرد؟

نکات:

- برای پیاده‌سازی از زبان برنامه‌نویسی Python و کتابخانه‌هایی مانند mlxtend یا apyori استفاده کنید.
- مستندات کد و تحلیل خود را در قالب یک گزارش ارائه دهید.
- ددلاین: ۱۴۰۴/۲/۱۷