



مهلت تحویل:

۱۴ اردیبهشت ۱۴۰۳

تمرین اول  
درس داده کاوی

۱. یکی از الگوریتم های Apriori و یا FP-Growth را به دلخواه خود برای مجموعه داده‌ی retail.dat (پیوست) پیاده سازی کنید. این مجموعه داده مربوط به تراکنش های مجموعه اقلام خریداری شده از یک خرده فروشی در بلژیک است. خروجی مطلوب الگوریتم شما، لیست مجموعه اقلام پرتکرار است.

لازم به ذکر است برای این تکلیف نیاز است که مقدار مناسب هایپرپارامتر  $\min\_sup$  توسط شما تعیین شود. همچنین در انتخاب زبان برنامه نویسی برای پیاده سازی این تکلیف مخیر هستید (البته مجاز به استفاده از کتابخانه های آماده این دو الگوریتم نمی باشید). چنانچه قواعد انجمنی حاصل از مجموعه اقلام پرتکرار را با استفاده از هایپرپارامتر  $\min\_conf$  مناسب به دست آورید، نمره اضافی به شما تعلق خواهد گرفت.

۲. الگوریتم های Apriori، FP-Growth و Eclat را از نظر مرتبه زمانی و فضایی در بهترین و بدترین حالت و همچنین حالت متوسط بررسی کنید و پاسخ خود را توضیح دهید.

برای سوال اول کد پیاده سازی شده به همراه خروجی الگوریتم خود را در یک پوشه قرار داده و همراه با پاسخ سوال دوم در یک فایل فشرده شده به نام خودتان به همراه شماره دانشجویی خود ارسال کنید (برای مثال Yourname\_9911111).

موفق باشید.

خوش خط و خوانا بنویسید و ترجیحا جواب ها تایپ شده باشد.