

به نام خدا

نام و نام خانوادگی : زهرا علیزاده

تاریخ : ۱۳۹۹/۷/۱۶

تمرین دوم (اختیاری): candidate_gen برای داده های table form در ARM به چه صورتی ست؟
استخراج کردن (mining) را می توان بر روی جداول رابطه ای نیز انجام داد. به این صورت که باید مجموعه داده به فرم جدول را به مجموعه داده تراکنشی تبدیل کرد که نسبتاً ساده و مناسب است اگر هر ویژگی یا attribute در جدول مقدار و ارزش categorical بگیرد. به سادگی هر مقدار به یک جفت attribute-value تغییر می کند که هر جفت attribute-value به عنوان یک آیتم در نظر گرفته می شود.
استفاده کردن از value در فرم تراکنشی کارا نیست چون ممکن است attribute های متفاوت value های یکسان یا مشترک داشته باشند.
برای یک مجموعه داده به فرم جدول، مرحله join یا ادغام متد candidate generation باید کمی اصلاح شود تا اطمینان حاصل شود که دو itemset را برای تولید یک candidate itemset شامل دو آیتم از یک attribute مشترک ادغام نمی کند.
همچنین می شود داده های تراکنشی را با استفاده از نمایش دودویی و برخورد با هر آیتم در I به عنوان یک attribute، به داده های table فرم تبدیل کرد. اگر یک تراکنش شامل یک آیتم باشد، مقدار attribute آن ۱ و در غیر این صورت ۰ است.