Sequential Pattern Mining

- Try to automatically find *temporal* patterns within the data set

ARM Example

- If person X buys diapers,
- Person X buys beer

- Purchases occur *at the same time*

SPM Example

- If person X takes Intro Stats now,
- Person X takes Advanced Data Mining in a later semester

- Conclusion: recommend Advanced Data Mining to students who have previously taken Intro Stats

- Doesn't matter if they take other courses in between

SPM Example

- Learners in virtual environments have different sequences of behavior depending on their degree of self-regulated learning
- High self-regulated learning: Tend to gather information and then immediately record it carefully
- Low self-regulated learning: Tend to gather more information without pausing to record it

(Sabourin, Mott, & Lester, 2011)

Different Constraints than ARM

- If-then elements do not need to occur in the same data point
- Instead
  - If-then elements should involve the same student (or other organizing variable, like teacher or school)
  - If elements can be within a certain time window of each other
  - Then element time should be within a certain window after if times

Sequential Pattern Mining

- Find all subsequences in data with high support

- Support calculated as number of sequences that contain subsequence, divided by total number of sequences

GSP (Generalized Sequential Pattern)

- Classic Algorithm for SPM
- (Srikant & Agrawal, 1996)

Data pre-processing

- Data transformed from individual actions to sequences by user

- Bob: {GAMING and BORED, OFF-TASK and BORED, ON-TASK and BORED, GAMING and BORED, GAMING and FRUSTRATED, ON-TASK and BORED}

Data pre-processing

- In some cases, time also included

- Bob: {GAMING and BORED 5:05:20, OFF-TASK and BORED 5:05:40, ON-TASK and BORED 5:06:00, GAMING and BORED 5:06:20, GAMING and FRUSTRATED 5:06:40, ON-TASK and BORED 5:07:00}

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：https://d.book118.com/018123074075006025