TER Master 24/25

MineMetrics: a Comparative Tool for Itemset Mining Libraries

Nadjib Lazaar

Professeur des universités
Université Paris-Saclay, Laboratoire LISN, CNRS
Équipe LaHDAK

lazaar@lisn.fr

MineMetrics Project

Comparative Tool

- Number of participants accepted in this TER: 10 students

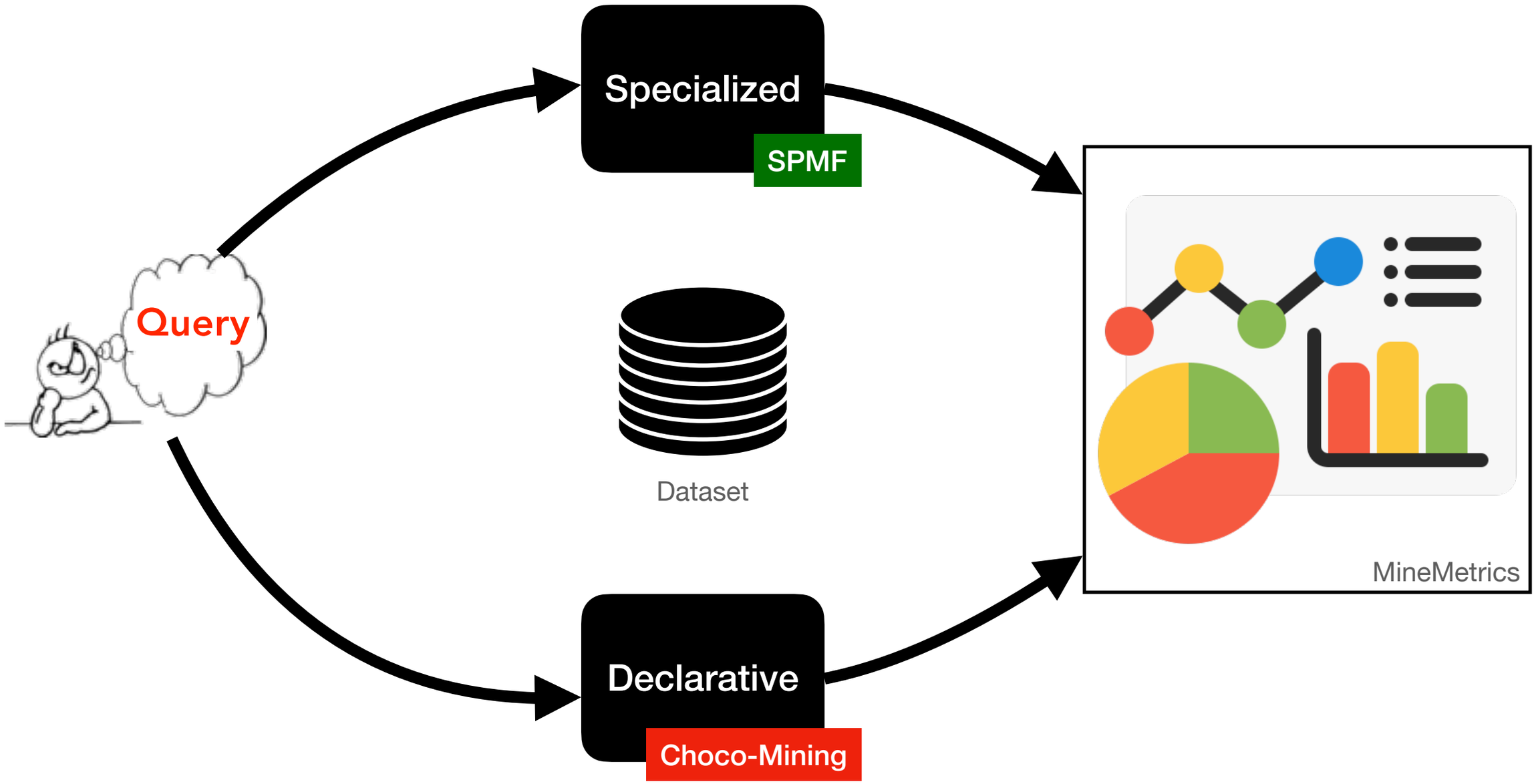
- The project aims to develop a tool for comparing two itemset mining libraries:
  - SPMF (specialized algorithms)
  - Choco-Mining (declarative method using Constraint Programming - CP)

- Comparaison based on performance metrics like CPU time, memory usage, and expressiveness

- Analyze results and draw conclusions about the strengths and weaknesses of each library based on experimental studies.

Itemset Mining

Specialized vs Declarative Methods



- SPMF: https://www.philippe-fournier-viger.com/spmf/
- choco-mining: https://gitlab.com/chaver/choco-mining
- FIMI datasets: http://fimi.uantwerpen.be/data/

MineMetrics Project

Requirements and Guidelines

- **Pre-requisites:** Knowledge in Data Mining; Basic programming skills (Java/Python); Front-end development basics (for visualization and user interface); Familiarity with performance evaluation metrics

- **Deliverables:**
  - **Final Report** (10-15 pages) including observations and conclusions from the comparative study
  - **Presentation** (15 minutes)

- **Weekly meeting** scheduled with students, which can be attended **remotely**.

- The project involves coding, testing, data analysis, front-end development, and drawing meaningful conclusions from the experimental comparison of the two tools.