A model to discover rare patterns using ML techniques

DBSE - TeamProject-2020
Supervisor - Sadeq Darrab
Team - 11
OVGU

Agenda

1. Motivation
2. Research Aim
3. Early literature review results
4. Tentative Timeline

Motivation

Fig 1

Fig 2

Rare Patterns & Challenges

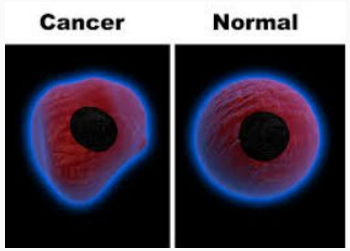


Fig 4

Rare patterns appears in less percentage of data

Hard to generalise

State of the art algorithms miss rare patterns because these rules have less than min support

If we lower the minsupport, number of rules generated explodes

So far >>>>>>>>>

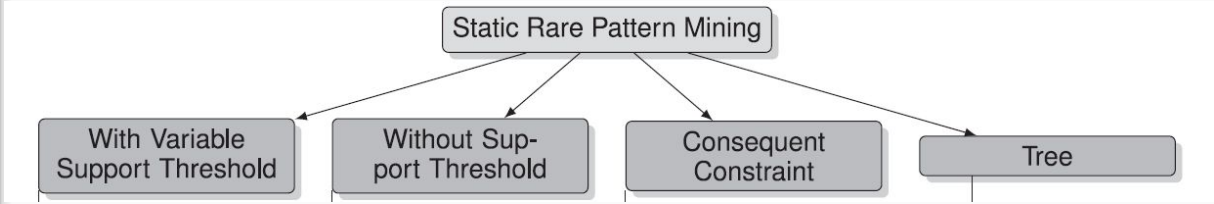
Why it doesn't work?

Almost all methods suffers from performance issues, no single method is guaranteed to work in all situations!

Streaming data

Noise or rare?

Scalability and performance



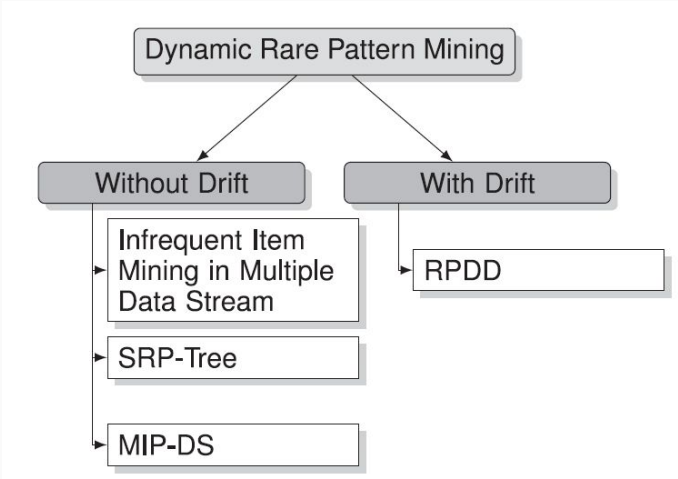
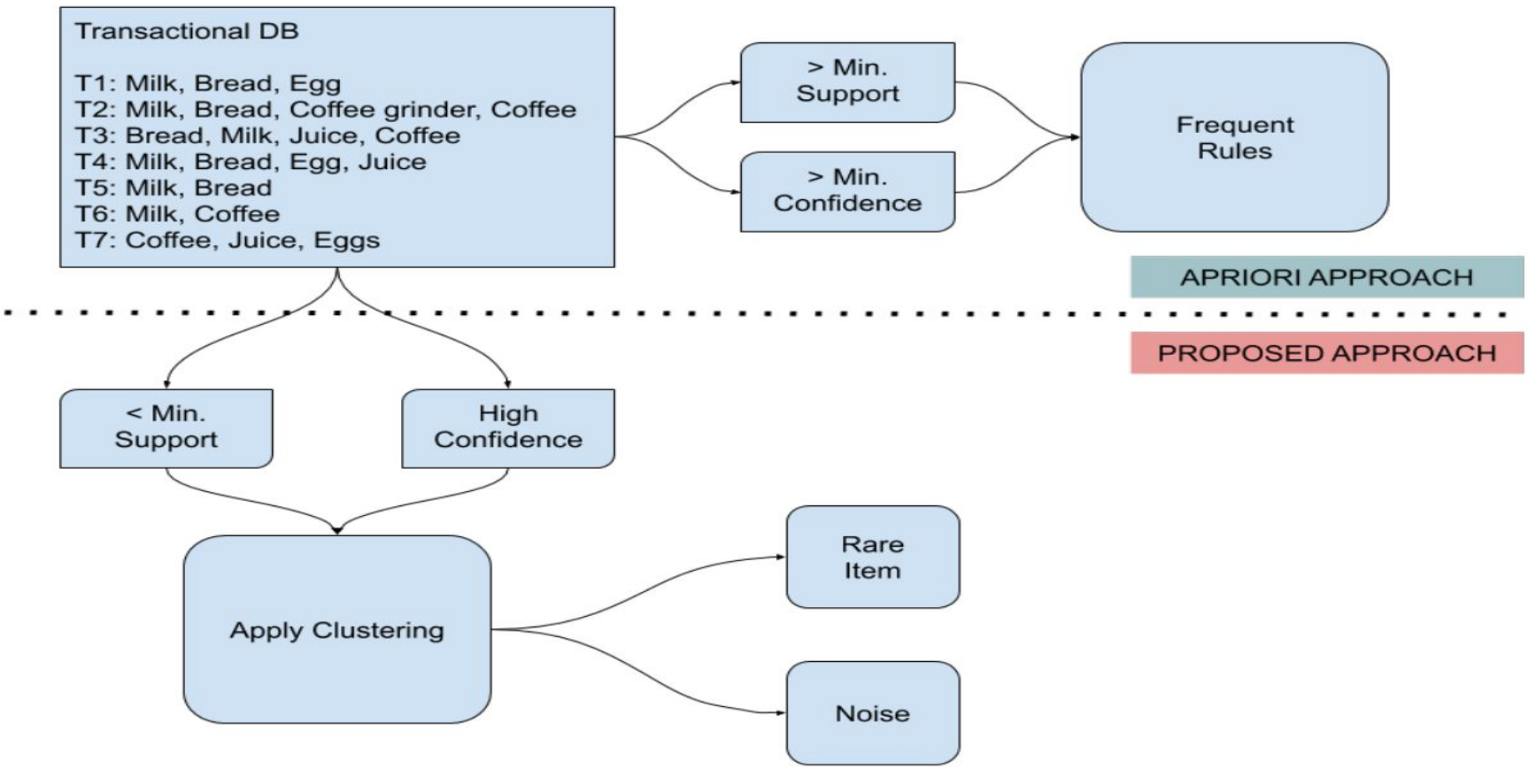
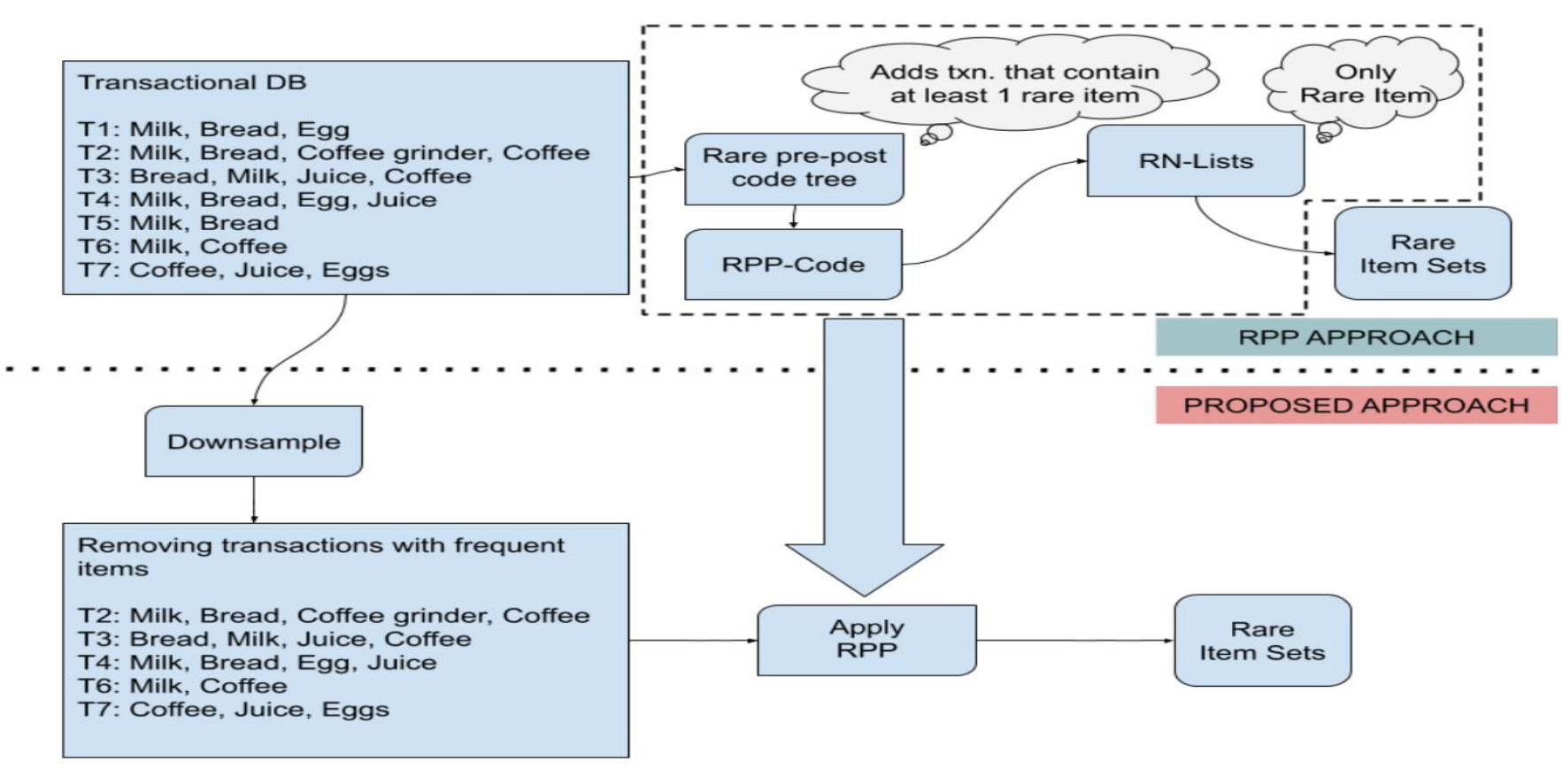


Fig 5,6

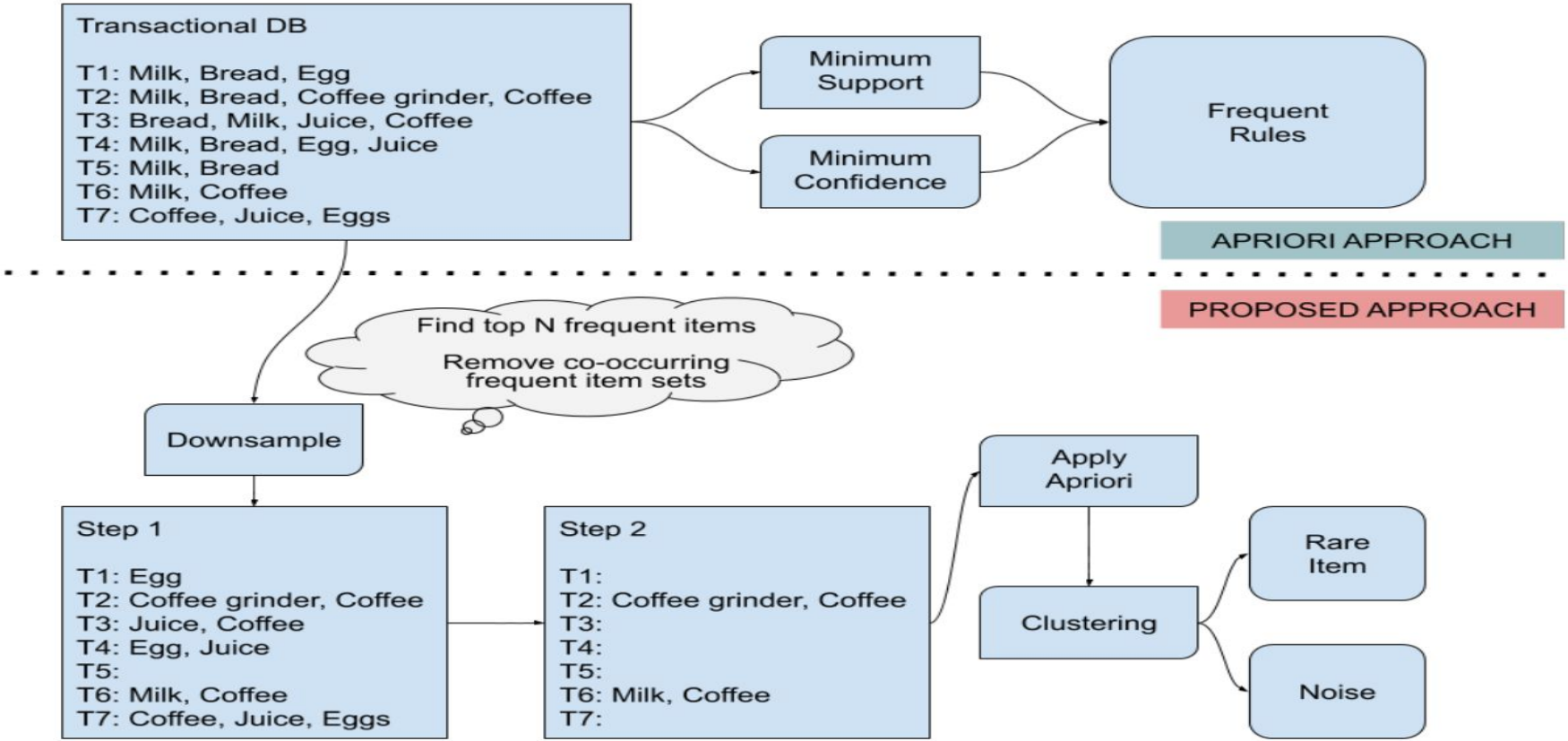
Approach 1- Consider data below min support with high confidence



Approach 2- First downsample then apply RPP



Approach 3 - Only downsampling + Apriori



Timeline


Project Timeline
Nov-20
Dec-20
Jan-21
Feb-21
TASK
W1
W2
W3
W4
W5
W6
W7
W8
W9
W10
W11
W13
W14
W15
Literature Review/ Current Workflow
Conceptualised design of proposed approach
Implementation of approach
Evaluation
Testing and Optimization
Documentation and Report Writing

Team

Madhu

Deeksha

Stanley

Seles

Surabhi

Priyam

Everyone gets a different role each sprint!

References-

Fig 1 to 4

www.google.com

Fig , 5,6

Yun Sing Koh and Sri Devi Ravana. 2016. Unsupervised rare pattern mining: A survey. ACM Trans. Knowl. Discov. Data 10, 4, Article 45 (May 2016), 29 pages.
DOI: http://dx.doi.org/10.1145/2898359

More to add...